

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/38, C12N 15/80, 1/15		A1	(11) International Publication Number: WO 98/01470
			(43) International Publication Date: 15 January 1998 (15.01.98)
(21) International Application Number: PCT/DK97/00305 (22) International Filing Date: 7 July 1997 (07.07.97) (30) Priority Data: 0740/96 5 July 1996 (05.07.96) DK (71) Applicant (for all designated States except US): NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK). (72) Inventor; and (75) Inventor/Applicant (for US only): CHRISTENSEN, Tove [DK/DK]; Novo Nordisk a/s, Novo Allé, DK-2880 Bagsværd (DK). (74) Common Representative: NOVO NORDISK A/S; Novo Allé, DK-2880 Bagsværd (DK).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(54) Title: A TRANSCRIPTION FACTOR (57) Abstract A transcription factor regulating $\alpha$ -amylase promoter initiated expression in filamentous fungi, especially in <i>Aspergillus</i> , DNA sequences coding for said factor, its transformation into and expression in fungal host organisms, and the use of said factor in such hosts for increasing the expression of a polypeptide of interest being produced by said host.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**Title: A Transcription Factor**FIELD OF THE INVENTION

5 The present invention relates to a transcription factor found in filamentous fungi, especially in *Aspergillii*, DNA sequences coding for said factor, its transformation into and expression in fungal host organisms, and the use of said factor in such  
10 hosts for increasing the expression of a polypeptide of interest being produced by said host.

BACKGROUND OF THE INVENTION

15 Transcription factors are well known proteins involved in the initiation of transcription. They have been studied intensively in many different organisms and have also been described in fungi. Dhawale and Lane (NAR (1993) 21 5537-5546) have recently compiled the transcription factors from fungi, including the  
20 filamentous fungi.

Many of the transcription factors are regulatory proteins; they bind to the promoter DNA and either activate or repress transcription as a response to stimuli to the cell.

25

The expression of the  $\alpha$ -amylase gene in *A. oryzae* is regulated in response to the available carbon source. The gene is expressed at its maximum when the organism is grown on starch or maltose (Lachmund et al. (1993) *Current Microbiology* 26 47-51;  
30 Tada et al. (1991) *Mol. Gen. Genet.* 229 301-306). The expression of  $\alpha$ -amylase is regulated at the transcriptional level as shown by Lachmund et al. (*supra*), which strongly suggests that transcription factors are involved in the regulation, but so far no gene for such a factor has been identified.

35

The promoter of the  $\alpha$ -amylase gene has been studied by deletion analysis (Tada et al. (1991) *Agric. Biol. Chem.* 55 1939-1941;

Tsuchiya et al. (1992) *Biosci. Biotech. Biochem.* 56 1849-1853; Nagata et al. (1993) *Mol. Gen. Genet.* 237 251-260). The authors of these papers propose that a specific sequence of the promoter is responsible for the maltose induction. Nagata et al. (*supra*)  
5 used this sequence as a probe in a gel shift experiment to see whether any proteins from *A. nidulans* nuclear extracts were able to bind to the promoter sequence. Three such proteins were found, but no involvement of these proteins in expression was shown. None of the proteins have been purified or identified by  
10 other means. Their genes likewise remain unknown.

#### SUMMARY OF THE INVENTION

The present invention relates to a transcription factor regulat-  
15 ing the expression of the  $\alpha$ -amylase promoter in filamentous fungi.

Accordingly, in a first aspect the invention relates to a DNA construct comprising a DNA sequence encoding a transcription  
20 factor of the invention, which DNA sequence comprises

- a) the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in *E. coli* ToC1058, DSM 10666, or
- b) an analogue of the DNA sequence defined in a), which
  - 25 i) is at least 60% homologous with the DNA sequence defined in a), or
  - ii) hybridizes with the same nucleotide probe as the DNA sequence defined in a), or
  - 30 iii) encodes a transcription factor which is at least 50% homologous with the transcription factor encoded by a DNA sequence comprising the DNA sequence defined in a), or
  - iv) encodes a transcription factor which is immunologically reactive with an antibody raised against the purified transcription factor encoded by the DNA sequence defined  
35 in a), or
  - v) complements the mutation in ToC879, i.e. enables ToC879 to grow on cyclodextrin and produce lipase when transformed with said DNA sequence.

The full length genomic DNA sequence encoding a transcription factor has been derived from a strain of the filamentous fungus *Aspergillus oryzae* and has been cloned into plasmid pToC320  
5 present in *E. coli* ToC1058, DSM 10666.

Said transcription factor encoding DNA sequence harboured in pToC320, DSM 10666, is believed to have the same sequence as that presented in SEQ ID NO: 1 and SEQ ID NO: 2. Accordingly,  
10 whenever reference is made to the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in DSM 10666 such reference is also intended to include the transcription factor encoding part of the DNA sequence presented in SEQ ID NO: 1 and SEQ ID NO: 2.

15 Accordingly, the terms "the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in DSM 10666" and "the transcription factor encoding part of the DNA sequence presented in SEQ ID NO: 1 and SEQ ID NO: 2" may be used  
20 interchangeably.

In further aspects the invention provides an expression vector harbouring the DNA construct of the invention, a cell comprising said DNA construct or said expression vector and a method of  
25 producing a peptide exhibiting transcription factor activity, which method comprises culturing said cell under conditions permitting the production of the transcription factor.

Such a transcription factor of the invention will typically  
30 originate from a filamentous fungus.

The term "filamentous fungus" is intended to include the groups Phycomycetes, Zygomycetes, Ascomycetes, Basidiomycetes and fungi imperfecti, including Hyphomycetes such as the genera  
35 *Aspergillus*, *Penicillium*, *Trichoderma*, *Fusarium* and *Humicola*.

The invention also relates to a method of producing a filamentous fungal host cell comprising the introduction of a DNA

fragment coding for any such factor into a filamentous fungus wherein an  $\alpha$ -amylase promoter or a co-regulated promoter regulates the expression of a polypeptide of interest in a manner whereby said factor will be expressed in said fungus.

5

In a further aspect the invention relates to a method of producing a polypeptide of interest, the expression of which is regulated by an  $\alpha$ -amylase promoter or a co-regulated promoter, comprising growing a filamentous fungal host cell as described  
10 above under conditions conducive to the production of said factor and said polypeptide of interest, and recovering said polypeptide of interest.

Finally the invention relates to the use of said factor for  
15 regulating the expression of a polypeptide of interest in a filamentous fungus.

In this context, regulation means to change the conditions under which the factor of the invention is active. This could mean  
20 different pH, substrate, etc. regimes, whereby the resulting effect is an improved regulation of the expression of the protein of interest.

Furthermore, regulation also comprises events occurring in the  
25 growth phase of the fungus during which the transcription factor is active. Depending on the circumstances, both advancing and/or postponing the phase wherein the factor is active may enhance the expression and thus the yield.

30 In addition, using standard procedures known in the art, the specific DNA sequences involved in the binding of a transcription factor may be identified, thereby making it possible to insert such sequences into other promoters not normally regulated by the factor and enabling those promoters to  
35 be under the regulation of said factor.

#### BRIEF DESCRIPTION OF THE TABLES AND DRAWING

In the figures

Fig. 1 shows the structure of the plasmid pMT1657, the construction of which is described in Example 1;

Fig. 2 shows the structure of the plasmid pToC316, the construction of which is described in Example 1;

10 Fig. 3 shows the structure of the plasmid pToC320, the construction of which is described in Example 1;

Fig. 4 shows the structure of the plasmids pToC342 and pToC359, the construction of which are described in Example 3;

15

Fig. 5 shows the structure of the plasmid pToC298, the construction of which is described in Example 4;

Fig. 6 shows the results of lipase production by a p960 transformant of *A. oryzae* IFO4177 cultured in YP media containing 2% glucose (—■—) or 10% glucose (—◆—), in comparison to ToC1075 cultured in YP media containing 2% glucose (—□—) or 10% glucose (—◇—) and described in Example 4;

25 Fig. 7 shows the results of lipase production by ToC1139 cultured in YP media containing 2% glucose (—■—) or 10% glucose (—◆—), in comparison to ToC1075 cultured in YP media containing 2% glucose (—□—) or 10% glucose (—◇—) and described in Example 4; and

30

Fig. 8 shows the autoradiograph results of *A. niger* DNA digested with the following restriction enzymes: lane 2, *XbaI*; lane 3, *XmaI*; lane 4, *SalI*; lane 5, *HindIII*; lane 6, *EcoRI*; lane 7, *BglIII*; lane 8, *BamHI*; lanes 1 and 9 contain <sup>32</sup>P-labelled 1 DNA digested with *BstEII*. The experiment is described in Example 5.

DETAILED DESCRIPTION OF THE INVENTION

In a first aspect the invention relates to a DNA construct comprising a DNA sequence encoding a transcription factor regulating an  $\alpha$ -amylase promoter, which DNA sequence comprises

- a) the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in *E. coli* ToC1058, DSM 10666, or
- b) an analogue of the DNA sequence defined in a), which
  - i) is at least 60% homologous with the DNA sequence defined in a), or
  - ii) hybridizes with the same nucleotide probe as the DNA sequence defined in a), or
  - iii) encodes a transcription factor which is at least 50% homologous with the transcription factor encoded by a DNA sequence comprising the DNA sequence defined in a), or
  - iv) encodes a transcription factor which is immunologically reactive with an antibody raised against the purified transcription factor encoded by the DNA sequence defined in a), or
  - v) complements the mutation in ToC879, i.e. enables ToC879 to grow on cyclodextrin and produce lipase when transformed with said DNA sequence.

As defined herein, a DNA sequence analogous to the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in *E. coli* ToC1058, DSM 10666, is intended to indicate any DNA sequence encoding a transcription factor regulating an  $\alpha$ -amylase promoter, which transcription factor has one or more of the properties cited under (i)-(v) above.

The analogous DNA sequence may be isolated from a strain of the filamentous fungus *A. oryzae* producing the transcription factor, or another or related organism and thus, e.g. be an allelic or species variant of the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in DSM 10666.

Alternatively, the analogous sequence may be constructed on the basis of the DNA sequence presented as the transcription factor encoding part of SEQ ID NO: 1 and SEQ ID NO: 2, e.g. be a sub-  
5 sequence thereof, and/or by introduction of nucleotide substitutions which do not give rise to another amino acid sequence of the transcription factor encoded by the DNA sequence, but which corresponds to the codon usage of the host organism intended for production of the transcription factor, or by  
10 introduction of nucleotide substitutions which may give rise to a different amino acid sequence.

When carrying out nucleotide substitutions, amino acid residue changes are preferably of a minor nature, that is conservative  
15 amino acid residue substitutions that do not significantly affect the folding or activity of the protein, small deletions, typically of one to about 30 amino acid residues; small amino- or carboxyl-terminal extensions.

20 Examples of conservative substitutions are within the group of basic amino acids (such as arginine, lysine, histidine), acidic amino acids (such as glutamic acid and aspartic acid), polar amino acids (such as glutamine and asparagine), hydrophobic amino acids (such as leucine, isoleucine, valine), aromatic  
25 amino acids (such as phenylalanine, tryptophan, tyrosine) and small amino acids (such as glycine, alanine, serine, threonine, methionine). For a general description of nucleotide substitution, see e.g. Ford, et al., (1991), Protein Expression and Purification 2, 95-107.

30 It will be apparent to persons skilled in the art that such substitutions can be made outside the regions critical to the function of the molecule and still result in an active transcription factor. Amino acid residues essential to the  
35 activity of the transcription factor encoded by a DNA construct of the invention, and therefore preferably not subject to substitution, may be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning

mutagenesis (cf. e.g. Cunningham and Wells, (1989), *Science* 244 1081-1085). In the latter technique mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for biological activity (i.e. transcription factor regulating an  $\alpha$ -amylase promoter) to identify amino acid residues that are critical to the activity of the molecule.

The homology referred to in (i) above is determined as the degree of identity between the two sequences indicating a derivation of the one sequence from the other. The homology may suitably be determined by means of computer programs known in the art such as GAP provided in the GCG program package (Needleman, S.B. and Wunsch, C.D., (1970), *Journal of Molecular Biology* 48 443-453). Using GAP with the following settings for DNA sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the DNA sequence exhibits a degree of identity preferably of at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95% with the transcription factor encoding part of the DNA sequence shown in SEQ ID NO: 1 and SEQ ID NO: 2.

The hybridization referred to in (ii) above is intended to indicate that the analogous DNA sequence hybridizes to the same probe as the DNA sequence encoding the transcription factor under certain specified conditions, which are described in detail in the Materials and Methods section hereinafter. The oligonucleotide probe to be used is the DNA sequence corresponding to the transcription factor encoding part of the DNA sequence shown in SEQ ID NO: 1 or SEQ ID NO: 2 or a fragment thereof.

The homology referred to in (iii) above is determined as the degree of identity between the two sequences indicating a derivation of the first sequence from the second. The homology may suitably be determined by means of computer programs known in the art such as GAP provided in the GCG program package (Needleman, S.B. and Wunsch, C.D., *supra*). Using GAP with the

following settings for transcription factor sequence comparison:  
GAP creation penalty of 3.0 and GAP extension penalty of 0.1,  
the transcription factor encoded by an analogous DNA sequence  
exhibits a degree of identity preferably of at least 50%, more  
5 preferably at least 60%, more preferably at least 70%, even more  
preferably at least 80%, especially at least 90% with the  
transcription factor encoded by a DNA construct comprising the  
transcription factor encoding part of the DNA sequence shown in  
SEQ ID NO: 2, e.g. with the amino acid sequence SEQ ID NO: 3.

10

In connection with property (iv) the immunological reactivity  
may be determined by the method described in the Materials and  
Methods section hereinafter.

15 In relation to the property (v) the complementation method is  
described in Example 1 herein.

The DNA sequence encoding a transcription factor of the  
invention can be isolated from the strain *Aspergillus oryzae* IFO  
20 4177 using standard methods e.g. as described by Sambrook, et  
al., (1989) Molecular Cloning: A Laboratory Manual. Cold Spring  
Harbor Lab.; Cold Spring Harbor, NY.

General RNA and DNA isolation methods are also disclosed in WO  
25 93/11249 and WO 94/14953, the contents of which are hereby  
incorporated by reference. A more detailed description of the  
complementation method is given in Example 1 herein.

Alternatively, the DNA encoding a transcription factor of the  
30 invention may, in accordance with well-known procedures, be  
conveniently isolated from a suitable source, such as any of the  
below mentioned organisms, by use of synthetic oligonucleotide  
probes prepared on the basis of a DNA sequence disclosed herein.  
For instance, a suitable oligonucleotide probe may be prepared  
35 on the basis of the transcription factor encoding part of the  
nucleotide sequences presented as SEQ ID NO: 1 or any suitable  
subsequence thereof, or on the basis of the amino acid sequence  
SEQ ID NO: 3.

The invention relates specifically to a transcription factor regulating the expression of the  $\alpha$ -amylase promoter in a filamentous fungus, which factor as indicated in Example 2 may even regulate the expression of other genes.

In this context the expression "filamentous fungus" is intended to include the groups Phycomycetes, Zygomycetes, Ascomycetes, Basidiomycetes and fungi imperfecti, including Hyphomycetes such as the genera *Aspergillus*, *Penicillium*, *Trichoderma*, *Fusarium* and *Humicola*.

In this context the expression " $\alpha$ -amylase promoter" means a sequence of bases immediately upstream from an  $\alpha$ -amylase gene which RNA polymerase recognises and binds to promoting transcription of the gene coding for the  $\alpha$ -amylase.

As indicated, transcription factors are known from many organisms and it is therefore expected that similar or corresponding factors may be found originating from other fungi of the genera *Aspergillus*, *Trichoderma*, *Penicillium*, *Fusarium*, *Humicola*, etc., having an enhancing effect on the expression of a polypeptide being under the regulation of amylase promoters in any fungus belonging to any of these genera.

A comparison of the DNA sequence coding for the transcription factor regulating the  $\alpha$ -amylase promoter has shown some degree of homology to a transcription factor (CASUCI) regulating the expression of glucosidase in *Candida* and to MAL63 of *Saccharomyces cerevisiae* as disclosed in Kelly and Kwon-Chung, (1992) *J. Bacteriol.* 174 222-232.

It is at present contemplated that a DNA sequence encoding a transcription factor homologous to the transcription factor of the invention, i.e. an analogous DNA sequence, may be obtained from other microorganisms. For instance, the DNA sequence may be derived by a similar screening of a cDNA library of another

11

microorganism, such as a strain of *Aspergillus*, *Saccharomyces*,  
*Erwinia*, *Fusarium* or *Trichoderma*.

An isolate of a strain of *A. oryzae* from which the gene coding  
5 for a transcription factor of the invention has been inactivated  
has been deposited by the inventors according to the Budapest  
Treaty on the International Recognition of the Deposit of  
Microorganisms for the Purposes of Patent Procedure at the DSM,  
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH,  
10 Mascheroder Weg 1b, D-38124 Braunschweig, DEUTSCHLAND.

Deposit date : 6 MAY 1996 (06.05.96)

Depositor's ref. : ToC879 = NN049238

DSM designation: *Aspergillus oryzae* DSM No.10671

15

The deposited strain *Aspergillus oryzae* DSM No.10671 can be used  
to isolate a transcription factor according to the invention  
from any strain of *Aspergillus oryzae* and any other fungal  
strain having such a gene by complementation as described  
20 hereinafter.

The expression plasmid pToC320 comprising the full length  
genomic DNA sequence encoding the transcription factor of the  
invention has been transformed into a strain of *E. coli*  
25 resulting in the strain ToC1058, which has been deposited by the  
inventors according to the Budapest Treaty on the International  
Recognition of the Deposit of Microorganisms for the Purposes of  
Patent Procedure at the DSMZ, Deutsche Sammlung von  
Mikroorganismen und Zellkulturen GmbH., Mascheroder Weg 1b, D-  
30 38124 Braunschweig, DEUTSCHLAND.

Deposit date : 6 MAY 1996 (06.05.96)

Depositor's ref. : ToC1058 = NN049237

DSM designation: *E. coli* DSM No.10666

35

According to the invention, factors of this type originating  
from the species *A. oryzae*, *A. niger*, *A. awamori*, etc.,  
especially *A. oryzae* IFO4177 are preferred.

The transcription factor of the invention has been found not only to be involved in the regulation of the  $\alpha$ -amylase promoter, but also in the regulation of the glucoamylase promoter from *A. oryzae*.

Especially, the invention comprises any factor having an amino acid sequence comprising one or more fragments or combinations of fragments of the amino acid sequence depicted as SEQ ID NO: 3.

Truncated forms of the transcription factor may also be active. By truncated forms are meant modifications of the transcription factor wherein N-terminal, C-terminal or one or more internal fragments have been deleted.

A further aspect of the invention relates to a DNA sequence coding for any of these factors.

In this aspect the invention especially comprises any DNA sequence coding for one or more fragments of the amino acid sequence depicted as SEQ ID NO: 3.

More specifically the invention relates to a DNA sequence comprising one or more fragments or a combination of fragments of the DNA sequence depicted as SEQ ID NO: 1 and SEQ ID NO: 2.

According to a further aspect the invention relates to a method of producing a filamentous fungal host cell comprising the introduction of any of the above mentioned DNA fragments into a filamentous fungus wherein the  $\alpha$ -amylase promoter or another co-regulated promoter regulates the expression of a polypeptide of interest in a manner whereby said factor will be expressed in said fungus.

The introduction of said DNA fragment may be performed by means of any well known standard method for the introduction of DNA

into a filamentous fungus, such as by use of an expression vector and host cells as described below.

Therefore, the invention also provides a recombinant expression  
5 vector comprising the DNA construct of the invention.

The expression vector of the invention may be any expression vector that is conveniently subjected to recombinant DNA procedures, and the choice of vector will often depend on the  
10 host cell into which it is to be introduced.

Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication,  
15 e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

20 In the expression vector, the DNA sequence encoding the transcription factor should either also contain the expression signal normally associated with the transcription factor or should be operably connected to a suitable promoter and terminator sequence. The promoter may be any DNA sequence which  
25 shows transcriptional activity in the host cell of choice and may be derived from genes that are either homologous or heterologous to the host cell. The procedures used to ligate the DNA sequences coding for the transcription factor, the promoter and the terminator, respectively, and to insert them into  
30 suitable vectors are well known to persons skilled in the art (cf., Sambrook, et al., supra).

Examples of suitable promoters for use in filamentous fungal host cells are, for instance, the *A. nidulans* ADH3 promoter  
35 (McKnight, et al. (1985) *The EMBO J.* 4 2093-2099) or the *tpiA* promoter. Examples of other useful promoters are those derived from the gene encoding *Aspergillus oryzae*  $\alpha$ -amylase, *Aspergillus niger* neutral  $\alpha$ -amylase, *Aspergillus niger* acid stable  $\alpha$ -

amylase, *Aspergillus niger*, *Aspergillus awamori*, or *Aspergillus*.  
or *oryzae* glucoamylase (*gluA*), *A. oryzae* alkaline protease (*alp*),  
*A. oryzae* nitrate reductase (*niaD*), *Aspergillus oryzae* triose  
phosphate isomerase (*tpi*), *Aspergillus nidulans* acetamidase, or  
5 an *Aspergillus* promoter coding for an amino acid biosynthetic  
gene such as *argB*.

In yet another aspect the invention provides a host cell  
comprising the DNA construct of the invention and/or the  
10 recombinant expression vector of the invention.

Preferably, the host cell of the invention is a eukaryotic cell,  
in particular a fungal cell such as a yeast or filamentous  
fungal cell. In particular, the cell may belong to a species of  
15 *Trichoderma*, preferably *Trichoderma harzianum* or *Trichoderma*  
*reesei*, or a species of *Aspergillus*, most preferably *Aspergillus*  
*oryzae* or *Aspergillus niger*. Fungal cells may be transformed by  
a process involving protoplast formation and transformation of  
the protoplasts followed by regeneration of the cell wall in a  
20 manner known per se. The use of *Aspergillus* as a host microorga-  
nism is described in EP 238 023 (Novo Nordisk A/S), the contents  
of which are hereby incorporated by reference. The host cell may  
also be a yeast cell, e.g. a strain of *Saccharomyces*, in  
particular *Saccharomyces cerevisiae*, *Saccharomyces kluyveri* or  
25 *Saccharomyces uvarum*, a strain of *Schizosaccharomyces* sp., such  
as *Schizosaccharomyces pombe*, a strain of *Hansenula* sp., *Pichia*  
sp., *Yarrowia* sp., such as *Yarrowia lipolytica*, or *Kluyveromyces*  
sp., such as *Kluyveromyces lactis*.

30 The endogenous *amyR* gene of the host cell may be deleted or  
inactivated by other means. The introduction of *amyR* control by  
a heterologous promoter will then lead to a completely new  
scheme of regulation of the  $\alpha$ -amylase promoter. If, for  
example, *amyR* is fused to the *A. oryzae niaD* promoter, the  $\alpha$ -  
35 amylase promoter will become inducible by nitrate. If, instead  
of the *niaD* promoter, a *palC*-regulated promoter is used, the  
activity of the  $\alpha$ -amylase promoter will be regulated by pH.

The invention also comprises a method of producing a polypeptide of interest, whereby a host cell as described above is grown under conditions conducive to the production of said factor and  
5 said polypeptide of interest, and said polypeptide of interest is recovered.

The medium used to culture the transformed host cells may be any conventional medium suitable for growing the host cells in  
10 question. The expressed polypeptide of interest may conveniently be secreted into the culture medium and may be recovered therefrom by well-known procedures including separating the cells from the medium by centrifugation or filtration, precipitating proteinaceous components of the medium by means of a salt such  
15 as ammonium sulphate, followed by chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

According to the invention the method may be used to produce a  
20 polypeptide of interest that is a medicinal polypeptide, especially such medicinal polypeptides as growth hormone, insulin, blood clotting factor, and the like.

The method of the invention may also be used for the production  
25 of industrial enzymes, such as proteases, lipases, amylases, glucoamylases, oxido reductases, carbohydrases, carbonyl hydrolases, cellulases, esterases, etc.

According to a further aspect of the invention said transcrip-  
30 tion factor may be used for enhancing the expression of a polypeptide of interest in a filamentous fungus, such as a fungus of the genus *Aspergillus*, *Trichoderma*, *Penicillium*, *Fusarium*, *Humicola*, etc., especially of the species *A. oryzae*, *A. niger*, *A. awamori*, etc., and specifically *A. oryzae*.

35

The transcription factor of the invention may thus be used to enhance the expression of a medicinal polypeptide, such as growth hormone, insulin, blood clotting factor, etc.

Also, the expression of industrial enzymes, such as proteases, lipases, amylases, glucoamylases, oxido reductases, carbohydrases, carbonyl hydrolases, cellulases, esterases, etc., may be enhanced by the use of the transcription factor of the invention.

The transcription factor may also be used to identify the sequences in the  $\alpha$ -amylase promoter to which it binds. For example, this could be done by making a GST-fusion protein with the DNA binding domain of AmyR, such as the zinc finger, for production in *E. coli*. Such fusion proteins may be conveniently made using commercially available kits, for example, "The GST Gene Fusion Kit" from Pharmacia. The purified GST-fusion protein can then be used in conventional *in vitro* techniques such as gel shift assays or DNA footprint analyses (Kulmburg, P., et al. (1992) *Molecular and Cellular Biology* 12 1932-1939; Lutfiyya, L.L., and Johnston, M. (1996) *Molecular and Cellular Biology* 16 4790-4797). The identification of the AmyR binding site will make it possible to insert these sequences in other promoters not normally regulated by AmyR.

#### MATERIALS AND METHODS

25

##### **Hybridization:**

Suitable hybridization conditions for determining hybridization between a nucleotide probe and an "analogous" DNA sequence of the invention may be defined as described below. The oligonucleotide probe to be used is the DNA sequence corresponding to the transcription factor encoding part of the DNA sequence shown in SEQ ID NO: 1, i.e. nucleotides 1691..2676 + 2743..3193 + 3278..3653 in SEQ ID NO: 1, or a fragment thereof, e.g. nucleotides 1770-1800 in SEQ ID NO: 1.

35

##### **Hybridization conditions**

Suitable conditions for determining hybridization between a nucleotide probe and a homologous DNA or RNA sequence involves

pre-soaking of the filter containing the DNA fragments or RNA to hybridize in 5x SSC (standard saline citrate buffer) for 10 min, and prehybridization of the filter in a solution of 5x SSC (Sambrook, et al., *supra*), 5x Denhardt's solution (Sambrook, et al., *supra*), 0.5 % SDS and 100 µg/ml of denatured sonicated salmon sperm DNA (Sambrook, et al., *supra*), followed by hybridization in the same solution containing a random-primed (Feinberg, A. P. and Vogelstein, B. (1983) *Anal. Biochem.* 132 6-13), <sup>32</sup>P-dATP-labeled (specific activity > 1 x 10<sup>9</sup> cpm/µg) probe for 12 hours at ca. 65°C. The filter is then washed two times for 30 minutes in 2x SSC, 0.5 % SDS at preferably not higher than 50°C, more preferably not higher than 55°C, more preferably not higher than 60°C, more preferably not higher than 65°C, even more preferably not higher than 70°C, especially not higher than 75°C.

Molecules to which the nucleotide probe hybridizes under these conditions are detected using a Phospho Image detector.

**20 Immunological cross-reactivity:**

Antibodies to be used in determining immunological cross-reactivity may be prepared by use of a purified transcription factor. More specifically, antiserum against the transcription factor of the invention may be raised by immunizing rabbits (or rodents) according to the procedure described by N. Axelsen et al. in: *A Manual of Quantitative Immuno-electrophoresis*, Blackwell Scientific Publications, 1973, Chapter 23, or A. Johnstone and R. Thorpe, *Immunochemistry in Practice*, Blackwell Scientific Publications, 1982 (more specifically pp. 27-31). Purified immunoglobulins may be obtained from the antisera, for example by salt precipitation ((NH<sub>4</sub>)<sub>2</sub> SO<sub>4</sub>), followed by dialysis and ion exchange chromatography, e.g. on DEAE-Sephadex. Immunochemical characterization of proteins may be done either by Ouchterlony double-diffusion analysis (O. Ouchterlony in: *Handbook of Experimental Immunology* (D.M. Weir, ed.), Blackwell Scientific Publications, 1967, pp. 655-706), by crossed immuno-electrophoresis (N. Axelsen et al., *supra*, Chapters 3 and

4), or by rocket immunoelectrophoresis (N. Axelsen et al., op cit., Chapter 2).

#### EXAMPLES

5

##### EXAMPLE 1

Cloning of the *amyR* transcription factor from *A. oryzae*

*amyR* was cloned by complementation of an *A. oryzae* mutant strain unable to express two different proteins both under control of the TAKA-amylase promoter. The mutant *A. oryzae* strain ToC879 was made by mutagenesis of a strain, SRe440, containing a lipase (HLL) encoding cDNA under control of the TAKA promoter and one copy of the TAKA-amylase gene transcribed from its own promoter.

15 The mutant was identified and isolated by its amylase negative (amylase<sup>-</sup>) phenotype and subsequently shown to be lipase negative (lipase<sup>-</sup>) as well.

The strain ToC879 contains intact copies of both expression cassettes. The amylase<sup>-</sup> phenotype makes ToC879 unable to grow on plates containing 1% cyclodextrin as the sole carbon source, while the parent strain SRe440 will grow on such plates.

ToC879 has been deposited at DSM under the name DSM No.10671.

25

*amyR* was isolated by co-transforming ToC879 with an *A. oryzae* cosmid library and an autonomously replicating pHelp1 based plasmid (D. Gems, I. L. Johnstone, and A. J. Clutterbuck (1991) Gene 98 61-67) carrying the *bar* gene from *Streptomyces* 30 *hygroscopicus* which confers resistance to glufosinate. The transformants were subjected to selection on plates containing cyclodextrin as the sole carbon source and screened for a concurrent reversion to the lipase<sup>-</sup> phenotype.

35 The transforming DNA was rescued from colonies able to grow on cyclodextrin. Subcloning resulted in the isolation of a 4.3 kb DNA fragment able to complement both phenotypes of ToC879. The gene harboured on this fragment was named *amyR*.

#### Construction of the pHelp1 derivative pMT1657

A plasmid, pMT1612, was made by ligation (and subsequent transformation into *E. coli* DH5a) of the following four fragments:

- 5 i) the *E. coli* vector pToC65 (described in EP 531 372) cut with *SphI/XbaI*,
- ii) a PCR fragment (containing the *A. nidulans amdS* promoter) cut with *SphI/BamHI*,
- iii) a 0.5 kb *BamHI/XhoI* fragment from pBP1T (B. Staubinger et al., (1992) *Fungal Genetics Newsletter* 39 82-83) containing the  
10 *bar* gene, and
- iv) a 0.7 kb *XhoI/XbaI* fragment from pIC AMG/Term (EP Application No. 87103806.3) containing the *A. niger* glucoamylase transcription terminator.

15

The PCR fragment containing the *amdS* promoter was made using the plasmid pMSX-6B1 (M. E. Katz et al., (1990) *Mol. Gen. Genet.* 220 373-376) as substrate DNA and the two oligonucleotides 4650 (SEQ ID NO: 4) and 4651 (SEQ ID NO: 5) as primers.

20

4650:	CTTGCATGCCGCCAGGACCGAGCAAG,	SEQ ID NO: 4
4651:	CTTGGATCCTCTGTGTTAGCTTATAG.	SEQ ID NO: 5

pMSX-6B1 contains an *amdS* promoter up mutation called I666.

25

pMT1612 was cut with *HindIII*, dephosphorylated and ligated to a 5.5 kb *HindIII* fragment from pHelp1 containing the *AMAl* sequence. The resulting plasmid, pMT1657 is self-replicating in *Aspergilli* and can be selected for by increased resistance to  
30 glufosinate. pMT1657 is depicted in Fig. 1, wherein Pamds represents the *amdS* promoter of fragment ii) above, *bar* represents fragment iii) above, and Tamg represents fragment iv) above.

#### 35 Construction of the cosmid library

A cosmid library of *Aspergillus oryzae* was constructed essentially according to the instructions from the supplier of the

"SuperCos1 cosmid vector kit" (Stratagene Cloning Systems, La Jolla CA, USA).

Genomic DNA of *A. oryzae* IFO4177 was prepared from protoplasts made by standard procedures (Christensen, T., et. al. (1988) *Biotechnology* 6 1419-1422).

After isolation the protoplasts were pelleted by centrifugation at 2500 rpm for 5 minutes in a Labofuge T (Heto); the pellet was then suspended in 10 mM NaCl, 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 µg/ml proteinase K and 0.5% SDS as stated in the manual from the Supercos 1 cosmid vector kit; the rest of the DNA preparation was done according to the instructions of the kit.

The size of the genomic DNA was analysed by electrophoresis using the CHEF-gel apparatus (Bio-Rad Laboratories, Hercules CA, USA). A 1% agarose gel was run for 20 hours at 200 volts with a 10-50 second pulse. The gel was stained with ethidium bromide and photographed. The DNA was 50-100 kb in size. The DNA was partially digested using *Sau3A*. The size of the digested DNA was 20-50 kb determined by the same type of CHEF-gel analysis as above. The CsCl gradient banded SuperCos1 vector was prepared according to the manual. Ligation and packaging were likewise performed as described in the manual.

After titration of the library, all of the packaging mix from one ligation and packaging was transfected into the host cells, XL1-Blue MR, and plated on 50 µg/ml ampicillin LB plates. Approximately 3800 colonies were obtained. Cosmid preparations from 10 colonies showed that they all had inserts of the expected size. The colonies were picked individually and inoculated in microtiter plate wells with 100 µl LB (100 µg/ml ampicillin) and incubated at 37°C overnight. 100 µl of 50% glycerol was added to each well, and the entire library was frozen at -80°C. A total of 3822 colonies were stored.

This represents the *A. oryzae* genome approximately 4.4 times. After picking the colonies the plates were scraped off, the scrape-off pooled and the total library was also stored in four pools as frozen glycerol stock. The four pools were named.  
5 ToC901-ToC904.

The individually frozen colonies in the library were inoculated onto LB-plates (100 µg/ml ampicillin) by using a multipin device of 6 rows of 8 pins fitting into half a microtiter dish. Plates  
10 were made containing colonies from all clones in the library.

The plates were incubated at 37°C overnight. Sterilized Whatman 540 filters cut to the size of a petri dish were placed upon the colonies which were incubated for two more hours at 37°C. The  
15 filters were transferred to LB plates containing 200 µg/ml of chloramphenicol and the plates were incubated overnight at 37°C.

The next day the filters were washed twice in 0.5 M NaOH for 5 minutes, then twice in 0.5 M Tris-HCl (pH7.4) for 5 minutes and  
20 then twice in 2x SSC for 5 minutes. The filters were wetted with ethanol and air dried.

#### Selection of *amyR* clones

Cosmid DNA was prepared from ToC901-904 and introduced into  
25 ToC879 by co-transformation with pMT1657. The transformation procedure is described in EP Application No. 87103806.3. Approximately 8700 transformants were selected by resistance to 1 mg/ml glufosinate in minimal plates (Cove D.J. (1966) *BBA* 113 51-56) containing 1 M sucrose for osmotic stabilization and 10  
30 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>.

Ten randomly chosen transformants were reisolated once on the same type of plates. Conidiospores from these 10 transformants were inoculated in minimal medium containing 1 mg/ml glufosinate  
35 and grown at 30°C until enough mycelium for DNA preparation could be harvested. DNA was prepared as described in T. Christensen et al. (*supra*).

The uncut DNA was applied to a 0.7% agarose gel, and electrophoresis was performed, followed by Southern blotting. The blot was hybridized with a <sup>32</sup>P-labelled SuperCos1 specific DNA fragment. Each of the ten transformants showed a band with a higher mobility than the linear chromosomal DNA. Each of the bands also hybridized to a pHelp1 specific probe, indicating that the co-transformation frequency of the cosmid library was close to 100% and that the cosmids had integrated into the autonomously replication vector pHelp1.

10

The transformants were unstable as expected for pHelp1 transformants. Less than 10% of the conidiospores from a glufosinate resistant colony gave rise to glufosinate-resistant progeny.

15 Conidiospores from all the transformants were collected in 8 pools and plated on minimal plates (Cove D.J., *supra*) containing 1 mg/ml glufosinate, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 1% b-cyclodextrin (Kleptose from Roquette Frères, 62136 Lestem, France)

20 Four colonies were obtained from one of the pools and one from one of the other pools. Two of the colonies from the first pool were reisolated once on the same kind of plates.

Conidiospores from the reisolated colonies were plated on  
25 minimal plates with either glucose or cyclodextrin as a carbon source and on glufosinate-containing plates. The glufosinate resistance and the ability to grow on cyclodextrin were both unstable phenotypes with the same degree of instability. This indicated that the gene conferring the ability to grow on  
30 cyclodextrin was physically linked to pMT1657 in the transformants.

Colonies from the reisolation plates were cut out and were analysed by rocket immune electrophoresis (RIE) using an  
35 antibody raised against the HLL lipase. The transformants gave a clear reaction with the antibody, while ToC879 colonies grown on maltose gave no reaction. This led to the conclusion that both the expression of amylase (i.e., growth on cyclodextrin) and

23

lipase (i.e. antibody binding) had been restored in these transformants. The gene responsible for this phenotype was named *amyR*.

#### 5 Isolation of the *amyR* gene

In order to rescue the *amyR* gene from the amylase<sup>+</sup>, lipase<sup>+</sup> transformants of ToC879, two different approaches were used successfully.

- 10 DNA was prepared from mycelium grown in minimal medium with cyclodextrin as the carbon source.

In the first approach the DNA was packaged into  $\lambda$ -heads using the Gigapack<sup>®</sup> II kit from Stratagene in an attempt to rescue the  
15 original cosmid out of the total DNA. The packaging reaction was incubated with XL1-Blue MR *E. coli* under the conditions specified by the kit supplier. The *E. coli* cells were plated on LB plates with 50  $\mu$ g/ml ampicillin. Two colonies appeared on the plates; the cosmids they contained were identical and named  
20 ToC1012.

In the second approach the total DNA was used in an attempt to transform competent *E. coli* DH5a cells. Sixteen colonies were isolated and shown to contain six different plasmids by  
25 restriction enzyme digest. Each of the plasmids was digested with *EcoRI* and subjected to Southern analysis. A <sup>32</sup>P-labelled probe of a mixture of pMT1657 and SuperCos1 was used to identify DNA fragments not part of any of these vectors. Two *EcoRI* fragments, approximately 0.7 and 1.2 kb in size, did not  
30 hybridize to any of these probes. The 1.2 kb fragment was isolated, labelled with <sup>32</sup>P and used as a probe in a hybridization experiment with the filters containing the part of the cosmid library that gave rise to the original transformants. Six cosmids from the pool (ToC904), containing approximately  
35 1000 clones did hybridize.

Of these, some were shown by restriction enzyme digestion to be identical, resulting in the isolation of four different cosmids.

All cosmids contained at least parts of the TAKA-amylase gene as well. The four cosmids and the cosmid ToC1012 were transformed into ToC879 by co-transformation with pMT1623, a pUC based plasmid that carries the *bar* gene under the control of the *A. oryzae tpi* promoter. Fifteen transformants from each co-transformation were isolated by resistance to glufosinate and tested for the ability to grow on cyclodextrin.

Eight transformants of ToC1012 and three transformants of one of the other cosmids, 41B12, were able to grow. None of the transformants of the other cosmids grew. That not all of the transformants of ToC1012 and 41B12 were able to grow is likely to be a reflection of the co-transformation frequency in each experiment. Colonies from the transformants growing on cyclodextrin were analysed by RIE, and showed that they all produced lipase.

DNA fragments obtained by digesting 41B12 with either *Bgl*II, *Hind*III or *Pst*I were cloned into pUC19 in order to subclone *amyR* from the cosmid. The subclones were transformed into ToC879 and the transformants analysed for the ability to grow on cyclodextrin and produce lipase as described above. As depicted in Fig. 2, one plasmid called pToC316 was shown to contain an approximate 9 kb *Hind*III fragment which was identified as containing *amyR*.

Further subcloning resulted in a plasmid called pToC320 containing a 4.3 kb *Hind*III/*Sac*I fragment, which is shown in Fig. 3 and was subsequently sequenced on an ABI DNA sequencer using both further subcloning and primer walking.

A DNA sequence of 3980 bp including the *amyR* gene is shown in SEQ ID NO: 1. The deduced amino acid sequence is shown in SEQ ID NO: 3 and reveals a Gal 4-type zinc finger sequence between amino acids 28-54. Such sequences are known to bind to DNA (Reece, R.J., and Ptashne, M. (1993) Science 261 909-910).

25

amyR maps close to one of the three amylase genes in IFO4177, since it was isolated from a cosmid also containing amylase-specific DNA fragments. Mapping of the cosmid showed that the  $\alpha$ -amylase gene and amyR are 5-6 kb apart. Southern analysis of genomic DNA showed that only one copy of amyR is present in IFO4177, and confirmed that it maps close to one of the amylase genes.

#### Analysis of amyR cDNA

10 mRNA was made by the method of Wahleithner, J. A., et al. (1996, Curr. Genet. 29 395-403) from a culture of *A. oryzae* grown in maltose containing medium under conditions favorable for  $\alpha$ -amylase production. Double stranded cDNA was made by standard procedures and used for PCR reactions with the following  
15 primers:

oligodT primer:	TTTGTAAAGCT <sub>31</sub>	SEQ ID NO. 9
23087:	CCCCAAGCTTCGCCGTCTGCGCTGCTGCCG	SEQ ID NO. 6
20865:	CGGAATTCATCAACCTCATCAACGTCCTTC	SEQ ID NO. 7
20 20866:	CGGAATTCATCGGCGAGATAGTATCCTAT	SEQ ID NO. 8

A PCR reaction with the primers 20866 and 23087 resulted in a fragment of approximately 1.1 kb. The fragment was digested with *EcoRI* and *HindIII*; these restriction sites were incorporated  
25 into the primers, and cloned into a pUC19 vector cut with the same enzymes.

The insert in the resulting plasmid was sequenced, the result located one intron in this part of the gene. The intron is  
30 indicated in SEQ ID NO: 2.

Another PCR reaction with the oligodT primer and primer 20866 did not result in a distinct fragment. An aliquot of this reaction was used as the starting point for a new reaction with  
35 the oligodT primer and the primer 20865, which resulted in a fragment of approximately 1.1 kb. This fragment was digested with *EcoRI* and *HindIII* and cloned into pUC19.

## EXAMPLE 3

## Overexpression of AmyR

A plasmid, pToC342, containing the coding region and 3' noncoding sequences of *amyR* fused to the promoter for the *A. oryzae tpi* gene was constructed. The *tpi* gene codes for triosephosphate isomerase, a constitutively expressed enzyme involved in primary metabolism. The *A. oryzae tpi* gene was isolated by crosshybridization with an *A. nidulans* cDNA clone according to the procedure of McKnight, G.L., et al, (1986, Cell 46 143-147). Sequencing led to identification of the structural gene. The promoter used was a fragment of approximately 700bp immediately upstream of the coding region. pToC342 was able to complement the mutation in ToC879. To pToC342 was further added the *A. oryzae pyrG* gene and the resulting plasmid, pToC359, was transformed into JaL250, a *pyrG* mutant of JaL228 described in patent application DK1024/96 filed 1996-09-19. Strains containing multiple copies of pToC359 were found to synthesise increased levels of glucoamylase.

20

## Construction of pToC342 and pToC359

A PCR reaction was made with pToC320 as the template and the following primers:

25 8753 GTTTCGAGTATGTGGATTCC  
8997 CGGAATTCCGATCCGAGCATGTCTCATTCTC

The resulting fragment was cut with *EcoRI*/*ApaI* to produce a fragment of approximately 180bp which was then cloned into pToC320 that had been digested with *EcoRI*/*ApaI*. The resulting plasmid, pToC336, was sequenced to confirm that the PCR fragment was intact. The 2.6kb *BamHI*/*SacI* fragment of pToC336 containing the coding region and the 3' untranslated sequence of *amyR* and an *EcoRI*/*BamHI* fragment of approximately 700bp containing the *tpi* promoter was cloned into *EcoRI*/*SacI* digested pUC19. The *BamHI* site downstream of the *tpi* promoter was introduced in vitro, whereas the *EcoRI* site is an endogenous site from the original *tpi* clone. The resulting plasmid, called pToC342, was

cut with *HindIII*, dephosphorylated and ligated to a 1.8 kb *HindIII* fragment containing the *A. oryzae pyrG* gene, resulting in a plasmid which was called pToC359. The structure of both pToC342 and pToC359 are shown in Fig. 4, wherein Ptpi represents the *tpi* promoter and TamyR represents the 3' noncoding region of *amyR*. The cloning of the *pyrG* gene has been previously described in WO 95/35385.

#### Expression in *A. oryzae* JaL250

JaL250 is a *pyrG* mutant of JaL228 selected by resistance to 5-fluoro-orotic acid. JaL228 has been described in patent application DK1024/96 filed 1996-09-19. JaL250 was transformed with pToC359 using standard procedures and by selecting for relief of uridine requirement. The transformants were reisolated twice through conidiospores and grown for four days in YP + 2% maltose at 30°C. Secreted glucoamylase was measured by the ability to cleave p-nitrophenyl  $\alpha$ -D-gluco-pyranoside. The transformants had 5-31 arbitrary glucoamylase units/ml in the fermentation broth, while JaL228 had 2-3 units/ml. The best transformant was named ToC1200. Southern analysis showed that multiple copies of pToC359 had integrated into the genome of ToC1200. Because of the  $\alpha$ -amylase promoter, ToC1200 may be used advantageously as a host strain for expression plasmids.

#### EXAMPLE 4

##### Carbon catabolite repression of the TAKA-promoter

The TAKA-amylase promoter is subject to carbon catabolite repression. In *Aspergilli* carbon catabolite repression is at least partially mediated via the transcriptional repressor CreA, a homologue to *S. cerevisiae* MIG1. The DNA binding sites in promoters under CreA control are known to be GC-rich and seemingly identical to the MIG1 sites in *S. cerevisiae*. The TAKA-amylase promoter contains several potential CreA binding sites. To determine whether this promoter is involved in carbon catabolite repression, three such sites were mutated, but provided only partial relief of carbon catabolite repression. In contrast, introduction of copies of constitutively expressed

AmyR in strains containing the modified promoter coupled to a reporter gene completely relieved repression of the reporter.

#### Construction of a CreA site deleted TAKA-amylase promoter

5 Three sites were identified as being potential CreA binding sites in the TAKA-amylase promoter by sequence comparison to known CreA and MIG1 sites. The resulting sites have the following sequences:

10 Site I      CCCCGGTATTG  
Site II      CCCCGGAGTCA  
Site III     ATATGGCCGGT

The bases underlined were changed to A's because such changes  
15 are known to destroy MIG1 binding sites. The substitutions were made using standard site-specific mutagenesis procedures. An expression vector, pToC297, containing the modified promoter and the 3' nontranscribed sequence of the glucoamylase gene from *A. niger* was constructed. pToC297 is identical to pToC68 described  
20 in WO 91/17243 except for the changes in the promoter. Both plasmids have a unique *Bam*HI site between the promoter and the terminator.

#### Expression of a lipase regulated by a CreA<sup>-</sup> TAKA-amylase 25 promoter

A *Bam*HI fragment of approximately 950bp containing the cDNA encoding a *Humicola lanuginosa* lipase was cloned into pToC297. (The cloning and expression of the *H. lanuginosa* lipase has been previously described in EP 305 216.) The resulting plasmid,  
30 pToC298, was transformed into *A. oryzae* IFO4177 by co-transformation with the *A. nidulans amdS* gene, and its structure is shown in Fig. 5, wherein Ptaka-creA represents the CreA binding site deficient TAKA-amylase promoter. The transformants were reisolated twice through conidiospores and one such  
35 transformant, ToC1075, which produces lipase, was chosen for further evaluation. ToC1075 and a p960 transformant of IFO4177 (previously described in EP 305 216) containing the lipase fused to the wild type TAKA-promoter were grown at 30°C in 10 ml YP

containing 2% or 10% glucose. Samples were taken daily for analysis of lipase in the fermentation broth. The lipase content was measured by rocket immune electrophoresis using a polyclonal antibody raised against purified lipase. Spent fermentation broth from *A. oryzae* IFO4177 did not react with the antibody. The glucose content of the fermentation broth was likewise measured daily using Tes-tape from Lilly.

On day one, glucose was detected in all cultures, but on day two glucose could be detected only in cultures originally containing 10%. The results of lipase production, shown in Fig. 6, indicate that the wild type promoter is repressed until glucose is no longer present. Thus, when the glucose becomes exhausted, lipase begins to accumulate. Fig. 6 also shows that the modified promoter is not as tightly regulated, as low levels of lipase are produced in the presence of glucose in the 10% glucose fermentation. Thus, there is partial glucose derepression seen in ToC1075.

#### 20 Relief of carbon catabolite repression of lipase in ToC1075 by pToC342

ToC1075 was transformed with pToC342 by co-transformation with the *bar*-containing plasmid, pMT1623. Strains containing multiple copies of pToC342 and which retained the lipase expression cassette were identified by Southern blot analysis; one such strain was. ToC1075 and ToC1139 were grown at 30°C in 10 ml YP containing either 2% or 10% glucose, and samples were assayed daily for lipase and glucose. The lipase was measured by cleavage of para-nitrophenyl-butyrate. The glucose content was measured with Tes-tape from Lilly. The results, shown in Fig. 7, indicate that ToC1075, as before, provides partial relief of glucose repression while lipase production by ToC1139 is independent of the presence of glucose.

#### 35 EXAMPLE 5

##### Southern analysis of *A. niger* for the *amyR* gene

The syntheses of  $\alpha$ -amylase and glucoamylase in *A. niger*, as in *A. oryzae*, are regulated by the carbon source. It is therefore

likely that *A. niger* also contains an *amyR* gene. This hypothesis was tested by looking for cross-hybridization between the *A. oryzae amyR* gene and *A. niger* chromosomal DNA.

5 DNA was prepared from *A. niger* by conventional methods. The DNA was cut with *Bam*HI, *Bgl*III, *Eco*RI, *Hind*III, *Sal*I, *Xma*I or *Xba*I, and the resulting DNA fragments were separated by electrophoresis on an agarose gel. The DNA was then blotted onto a nitrocellulose membrane and hybridized with a <sup>32</sup>P-  
10 labelled probe containing part of the structural *A. oryzae amyR* gene. The probe was made by PCR on pToC320 and starts at bp. no. 1683 and ends at bp. no. 2615 as shown in SEQ ID NO: 1. The hybridization was conducted in 10x Denhardt's solution, 5x SSC, 10mM EDTA, 1% SDS, 0.15 mg/ml polyA, 0.05 mg/ml yeast tRNA) at  
15 50°C overnight. After hybridization the membrane was washed under conditions of increasing stringency and the radioactivity on the membrane analysed by a PhosphoImager. Figure 8 shows the result when the membrane had been washed in 2x SSC, 0.1%SDS at 58°C. Unique bands can be seen with several of the restriction  
20 enzymes. Thus, the *A. niger amyR* gene can be cloned on the basis of this cross-hybridization result.

#### REFERENCES CITED IN THE SPECIFICATION

25

##### Description:

Dhawale and Lane (1993) NAR 21 5537-5546

Lachmund et al. (1993) Current Microbiology 26 47-51

30

Tada et al. (1991) Mol. Gen. Genet. 229 301-306

Tada et al. (1991) Agric. Biol. Chem. 55 1939-1941

35 Tsuchiya et al. (1992) Biosci. Biotech. Biochem. 56 1849-1853,

Nagata et al. (1993) Mol. Gen. Genet. 237 251-260

Ford et al., (1991), Protein Expression and Purification 2, 95-107

Cunningham and Wells, (1989) Science 244 1081-1085

5

Needleman, S.B. and Wunsch, C.D., (1970) Journal of Molecular Biology 48 443-453

Sambrook, J., et al., (1989), Molecular Cloning: A Laboratory  
10 Manual. Cold Spring Harbor Lab.; Cold Spring Harbor, NY.

WO 93/11249

WO 94/14953

15

Kelly and Kwon-Chung (1992) J. Bacteriol. 174 222-232

McKnight et al. (1985) The EMBO J. 4 2093-2099

20 EP 238 023

Kulmburg, P., et al. (1992) Molecular and Cellular Biology 12  
1932-1939

25 Lutfiyya, L.L., and Johnston, M. (1996) Molecular and Cellular  
Biology 16 4790-4797

Examples:

Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. 132:6-  
30 13

N. Axelsen et al. in: A Manual of Quantitative Immuno-electro-  
phoresis, Blackwell Scientific Publications, 1973, Chapter 23

35 A. Johnstone and R. Thorpe, Immunochemistry in Practice,  
Blackwell Scientific Publications, 1982, pp. 27-31

O. Ouchterlony in: Handbook of Experimental Immunology (D.M. Weir, ed.), Blackwell Scientific Publications, 1967, pp. 655-706

D. Gems, I. L. Johnstone, and A. J. Clutterbuck, (1991) *Gene* 98  
5 61-67

EP 531 372

B. Staubinger et al. (1992) *Fungal Genetics Newsletter* 39 82-83  
10

M.E. Katz et al. (1990) *Mol. Gen. Genet.* 220 373-376

Christensen, T., et. al., (1988) *Biotechnology* 6 1419-1422

15 Cove D.J., *BBA* (1966) 113 51-56

Reece, R.J., and Ptashne, M. (1993) *Science* 261 909-910

Wahleithner, J. A., et al. (1996) *Curr. Genet.* 29 395-403  
20

Hata, Y., et al. (1992) *Curr. Genet.* 22 85-91

McKnight, G.L., et al, (1986, *Cell* 46 143-147)

25 DK1024/96

WO 95/35385

WO 91/17243

30

EP 305 216

## SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

(i) APPLICANT:

10 (A) NAME: Novo Nordisk A/S  
(B) STREET: Novo Alle  
(C) CITY: Bagsvaerd  
(E) COUNTRY: Denmark  
(F) POSTAL CODE (ZIP): DK-2880  
(G) TELEPHONE: +45 4442 2668  
15 (H) TELEFAX: +45 4442 6080

(ii) TITLE OF INVENTION: A transcription factor

(iii) NUMBER OF SEQUENCES: 9

20 (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 3980 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus oryzae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

45 TCTAGACGGG CCACTGTGGG GTGCGGCAAG TTGATGCGG ACGGTGTGT AGTTGCTTCT 60  
TTTAAAGAAC GGCACCGCTC TGGGCTCTCC GAACCGGAT TGTAGCTAGA TGATATGTC 120  
50 TTGACGAACC AGGIGTCAC GGCACAATCC CTCACAATG ATGGCGCGTC CGTTTCCCAT 180  
CGATTGTGTC TACCTGGCGT GCAAGGCRAA ACATCGCGT CAACGTGCG AGGGGCATTG 240  
CCTGCAATCT CTGACCATG AGAGGGGAAG CAAGTCACG TAGTTGCRAG GGTATAGGTC 300  
55 CTAGCGAGCA ATGAGGTGGC TTCACCGTA CCGAGTGGG ACAGCATGAT CAGCGCTTTT 360  
GGGAACGTGA CGAAGAGTA CCGGTACGC GGAAGATGG AGATGAATCT CTGCGAGCA 420  
60 AAGGACGAGA CCGGAAAGA GTGTGTGAT TCTTGGGAG AGTACAGTA CTTGCGTTC 480  
CGGAATGG AAAAGTCTT GAACATGCT GCGATCATC TGATATCGT ACGCTGATTG 540  
GTGCATCGCC CGATAAATG CCGACAGAC GCTTGAAGCC TGAAAAGTA GTATTCTG 600  
65 AGAGATCCAT TCACAGAGT CAATCTGCG AAATACATG TTCCCGACT CATATTCCAA 660  
GGTGCGTAAA CCGCTCGGT GTGCGGTA GGGTTTCCA CGCATCTCT AGTGGTCCA 720

	TGA000GAGC ATCCGATGGC TTCCAGTATT GGGTGGTTGG GATGGACAAC AAGCTCCAAA	780
	TAAG00GAAT TTG0CTTTGG TCCAGGAATG AAGTCC00GT GGGGACCAGC GGCTCAG00C	840
5	AGGCTAAGAG TCGAATATCG TCATAGACCT TGGGCTCATG GGAGGTT0GG AGGTGTTA0G	900
	AT0CTCITCA ATG0CATICA TTCTCTGTTT TGA0CT0GGC TT00GAGAG TGGT00CT0C	960
10	CTTACAT00C CACATGCTGG ATGCAAG0CT GTGGTAC0CT GTTCTTTTCA GAAGTAGCAG	1020
	GCTAGGTICA CGATGAGCTG CCTTICAAAC CTGGAATAAC CATTACGTGA GACTGTCTTA	1080
	CTTCTTGAAT TGAT00CTGA CTAGAGTCTG CTCTAATAIG CTGTGTGGCA GGG00GGTCC	1140
15	CCT0GGGGIT GCTAAG0CTG ATTATG0CAC T00GTACAGT ATAAC00CAGG GT0GCTATAG	1200
	ATT00CTGCA TCTT00AG0C T00CTCAGAA CCTGATT0CA CCATTCTTAA G0GG00GTTA	1260
	G0CT0GAT0G GGTATAATGG AGTTA0CTAT AAACAG0ACT CTACAA0GAA T000GATGIG	1320
20	AGTTT0GAAC GAGTTGTAC G0ATGG0T0C T00CATTTGT TAGGAGTGAC GCTAG00CAC	1380
	CTTTAGG0CA CAGACTAAAC CAAGACAAAG ATGGAGTAGA CT0CAGGTAG ATTAAIT0CA	1440
25	ATCTCTT0CC CAAAGTAA0G G0GGGTTTTT TGCA0CT0CA G0CTCTTTTT TTCTTTTTTT	1500
	CTTTTTTTTT TTTTTTATTT GTT000CAGA TTCTTTTTCT TTTTCTTCAA T0CTGACGTT	1560
	CTCAAC00TG ATGG0GACAC AG0000CTTC GCTAT00CTC GCTTTTAC0GT C0G0CATCTC	1620
30	TCTAGTTGCT CT0G0GGAT G0CATGATTT CTAAG00CTC CACAT000G AGATAGTATC	1680
	CTAT00GAGC ATGTCTCATT CT0CAAC0GA CAT00CTCA ACAT00GAAA AGGAAAT0GA	1740
35	GTCAAC00CA GAAAAG00GC CTAACAG0C CT00GACAAT TG00GT0GAC GCAAAATCA	1800
	GTGTCTTAGA GAGCTTCCAT G0GACAAGTG C0AG0GTCTT CTCTCT0CT GTT0CTACAG	1860
	C0AG0T0CTC GGT00CAAGG G000CAAGTT C0GCA00CTC TAC0CTCT0G CT00CAT0CA	1920
40	T0CACT00CC T0CAGAC0AC GTCTCT0CAC CAAGGAATGG CTG00000AA A00CAGGG0C	1980
	TGT0CATTTG G0GT0000GA CGTCT000C GT0CAC0GTA G0GGA000C AGTATCTACA	2040
45	T0CAGACTTC T0GAGT0GT T0ACT0GACT A0C00000CA GATCT0GTCT OCTCT000GA	2100
	CT0GACAAAC T0GCTAT0G ACT0GT0CAC TAT0000GCA CT000000C CAC000GTCT	2160
	GT0GAG0CA AACCTTCTAG C0CATGTCAA TGTCTT0CTC AAGTAC0TGT T00GATCAT	2220
50	G000GT0GTG AGACAGGACC AGCT0CAGCA G0ACTG0CAC CAG000GAGC GCTTGTCT0C	2280
	C0AAC0CTAC GCTTTCATTG C0GCTCTATG C00000CAG CACAT0CAAC TGAAGCTGGA	2340
55	C0GT0CAGCA C0GGGT000G AG0000CTTC C0000GAG0C AGCCT0GAG GACAT0CTAT	2400
	GTGT000GA GAAGA0CT0C T0GCTGAAGC GGT0000GCA AGAAAGGAAT ACA0GTGGT	2460
60	CGA0GAAAT AACAT0GAAA A0CT0CTAAC CT0CTCTTT CTCTT000G OCTAC0GAAA	2520
	CCTAGACAGA CAGGATCAGS C0TGGTCTTA OCTATGT0AG A0CAGT0CA T0GTCTCAC	2580
	ACTAG00CTA CA000GAAT CCACATCTC GAACT0AGC GT0GAGGAAG CAGAAGAGAA	2640
65	AAGGAGAGTA TTCTGGCTCT TATT0GTAC AGAAAGGTAA GAAAAGAAAA AACTCTACTT	2700
	T00CAATCAC CACCA0GTAC CAAAAT0AC A0GAAAA0C AGAGGCTAGS CATTACAACA	2760
	AGCAAA0CA GTCATGCT0C GCACT0CAT CCACAA0CA CAGGT0CTGT GCTCAGACGA	2820

	CCCAATCCCTA GCTTACGGGT TCATCAACT CATCAAGTC TTGAAAAGC TCAGCCCAA	2880
5	TTCTTAAGAC TGGGTCTCG OGGGGGAG CAGGGGAG GGGACCCC GGGTACTTC	2940
	TTCTATCAA TCCAGTCTG CCAAGCAAT CTCCCTGAG GGGTCTCG AGATCAGAA	3000
	AGTAGACATC CTCATCTC AGCAATGGT ACRAACATG ATGTGAAAC TCTCATGAC	3060
10	CCAAGTCACA CAGCCCCCT CTGGGATGA GGGGTCTC CCTTCCAC TCCCGTGT	3120
	AGTGGGCAAG GGGTCTCG GGGTCTCG GGGGATC CAAGGTCTG TTGAGCTCA	3180
15	TGGTATCGA ATGGTAAGAA AGGAGCTTA CTTATCACA CCTCCCTCA TCAGTCTC	3240
	CCCATCTCT ATACCCGCA TCTACAAA ACGGAGGA CAAACTCT AGGAGCTGG	3300
	CACTCCGTA GGGAGTCT CCGCTCCCT AAGCAAAA GGGGCCCC ACTCGCGA	3360
20	ATGAGCATC GAGCCCCGAG AACTCTCTG GGGATCTC ACAACCTAT CCGGATCG	3420
	CGGTCCAA TCACTCTT TCCAGGCT CTGGAGCA AGTGGGCA TCACTGTT	3480
25	CGACTGTG CTTCCATCA GTGACTTCT GCTTGTCT GGTGGGCG CCGCTATT	3540
	GTTGGGAG GGTGAATCT GGTGTGTTT ATTGGGATC GGGATGATT TGCAAGAG	3600
	GGAGATGAG GGTGGGAG GGTGTGTT GGTGGGAG GAGATTGTT TTGAGGGG	3660
30	CTCTTTCTT TTCTTTGT GGTGTGTT GTGGGTTT TCTGGGGG GGGGGTGA	3720
	TATAGCTTG ACGATGCA TTGGGATGG GGTCTCTT GGTATATAT ATGGATTGT	3780
35	TTGTATAG TCCCTGGAG ACGTGAAT GATGTGGGA TCAATCTT CTAGGACT	3840
	GGAGCAGG GTTGGTTC TCGGTATT CTGATGTA GATTATAG AATCAGTAA	3900
	TGATCATAT TGTACATC TTAAGAAG ATATGCTTG CACCCGATA TGCAATAGA	3960
40	AACTGGTCT TCATTAGA	3980

## (2) INFORMATION FOR SEQ ID NO: 2:

- |    |   |
|----|---|
| 45 | (i) SEQUENCE CHARACTERISTICS:           |
|    | (A) LENGTH: 3980 base pairs             |
|    | (B) TYPE: nucleic acid                  |
|    | (C) STRANDEDNESS: single                |
|    | (D) TOPOLOGY: linear                    |
| 50 | (ii) MOLECULE TYPE: DNA (genomic)       |
|    | (iii) HYPOTHETICAL: NO                  |
| 55 | (iv) ANTI-SENSE: NO                     |
|    | (vi) ORIGINAL SOURCE:                   |
|    | (A) ORGANISM: <i>Aspergillus oryzae</i> |
| 60 | (ix) FEATURE:                           |
|    | (A) NAME/KEY: exon                      |
|    | (B) LOCATION: 1691..2676                |
| 65 | (ix) FEATURE:                           |
|    | (A) NAME/KEY: intron                    |
|    | (B) LOCATION: 2677..2742                |
|    | (ix) FEATURE:                           |
|    | (A) NAME/KEY: exon                      |
|    | (B) LOCATION: 2743..3193                |

(ix) FEATURE:  
 (A) NAME/KEY: intron  
 (B) LOCATION:3194..3277

5 (ix) FEATURE:  
 (A) NAME/KEY: exon  
 (B) LOCATION:3278..3653

10 (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION:join(1691..2676, 2743..3193, 3278..3653)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

15 TCTAGACCGG CCATGTGTGT GTCCGCCAAG TTGATTCCCG ACCGTGTGT AGTTGCTTCT 60  
 TTAAAGAAC GGCACCCCTC TCCCGTCTCC GAACCGAAT TGTAGCTAGA TGTATATGTC 120  
 20 TTGACGAACC AGGTGTCCAC GGCCAATCC CTCACAATTG ATGCCCCTGC CGTTTCCCAT 180  
 CGATTGTGTC TACCTGCGGT GCAAGGCAAA ACATCCCGT CAAAGTCCG AGGGGCATTG 240  
 CCTGCAATCT CTGACCATG AGAGGGGAG CAAGTCAAC TAGTTGCAG GGTATAGGTC 300  
 25 CTACGCAGCA ATGAGGTGGC TTCACCGTA CGGAGTGGG ACAGCATGAT CAAGCCTTTT 360  
 GCGAAGTGA CGAAGAGTA CCGTAAAGC GAGGATGG AGATGAATCT CTGCGAGCA 420  
 30 AAGGACGAGA CCGAAAGA GTGTGTGAT TCTTGGAGC AGTTACAGTA CTTCGTGTC 480  
 CGGAATTTG AAGGTTCCT GACCAATGCT GCGATCATC TGATATCCCT ACCCTGATTG 540  
 GTCCATCCCC CGATAAATGC CGACACGAC GCTTGAAGCC TGAAAAGTA GTATTCTCG 600  
 35 AGAGATCCAT TCACAGAGT CAATACGTC AATACATCG TTCCCACT CATATTCCAA 660  
 GGTGCTAAA CCCCCTCGT GTGCGGTA GGTTTTCA GGCATCTCT AGTGTGCCA 720  
 40 TGACGGGAGC ATCGATGGC TTCCAGTAT GGTGTGTG GATGACAAC AAGCTCCAA 780  
 TAAGGGGAT TTGCTTTG TCCAGGAATG AAGTCCCGT GGGACGAC GCTCAGGCC 840  
 AGGCTAAGAG TGAATATCG TCATAGACT TGGCTCATG GGAGTTGG AGGTGTACG 900  
 45 ATCTCTTCA ATGCCATCA TTCTGTGT TGACTGGC TTCCGAGAG TGGTGCTCC 960  
 CTACATCC CACATGCTG ATGCAAGCT GTGTAAGCT GTTCTTTCA GAAGTAGCAG 1020  
 50 GCTAGTTCA GATGAGCTG CCTTCAAC CTGGAATAC CATTAAGTGA GACTGTCTA 1080  
 CTCTTGAAT TGATCCCTGA CTAGAGTCT CTCATATG CTGTGGCA CGCGGCTCC 1140  
 CTTGGGGIT GCTAAGGCTG ATTTATGAC TCGTACAGT ATAACCCGG GTGGCTATAG 1200  
 55 ATTCCCTGCA TCTTCCAGC TCCCTACAA CCGATTCA CCATCTTAA GCGCGGTA 1260  
 GCTCGATGG GGTATATGG AGTTAACTAT AAACAGACT CTACACGAA TCCGATGTG 1320  
 60 AGTTTGAAC GAGTTGTAC GATGGGTC TCCATTGT TAGGATGAC GCTAGGGAC 1380  
 CTTTAGGCA CAGACTAAC CAAGACAAAG ATGAGTAGA CTCAGGTAG ATTAATCCA 1440  
 ATCTCTGTC CAAAGTAAAG CGGGTTTTT TGACCTGCA GCTCTTTTT TTCTTTTTT 1500  
 65 CTTTTTTTC TTTTTTAT GTTCCCGA TTCTTTCT TTTCTTCA TCTGAGTT 1560  
 CTCACCGTG ATGGGACAC AGCCCTTC GCTATCCCT GCTTTAGST CGGCATCT 1620

38

	TCTAGTTGCT CTGGGGGAT GCCATGATTT CTAAAGGCTC CACATGGGG AGATAGTATC	1680
5	CTTATCGAGC ATG TCT CAT TCT CCA ACC GAC ATT CCC TCA ACA TOC GAA Met Ser His Ser Pro Thr Asp Ile Pro Ser Thr Ser Glu 1 5 10	1729
10	AAG GAA ATG GAG TCA ACC CCA GAA AAG CCG CCT AAA CAG GGC TGC GAC Lys Glu Met Glu Ser Thr Pro Glu Lys Pro Pro Lys Gln Ala Cys Asp 15 20 25	1777
15	AAT TGC CGT CGA CGC AAA ATC AAG TGT TCT AGA GAG CTT CCA TGC GAC Asn Cys Arg Arg Arg Lys Ile Lys Cys Ser Arg Glu Leu Pro Cys Asp 30 35 40 45	1825
20	AAG TGC CAG CGT CTT CTT CTC TCC TGT TCC TAC AGC GAC GTG CTC CGT Lys Cys Gln Arg Leu Leu Leu Ser Cys Ser Tyr Ser Asp Val Leu Arg 50 55 60	1873
25	CGC AAG GGC CCC AAG TTC CGC AGC CTC TAC CCT CTC GCT CCC ATC CAT Arg Lys Gly Pro Lys Phe Arg Thr Leu Tyr Pro Leu Ala Pro Ile His 65 70 75	1921
30	CCA CTC GGC TCA CGA CCA CGT CCT CTC ACC AAG GAA TGG CTG CCC CCA Pro Leu Ala Ser Arg Pro Arg Pro Leu Thr Lys Glu Trp Leu Pro Pro 80 85 90	1969
35	AAC CCA GGG GCT TGC CAT TTG GCG TCC CCG AGC TCT CCG CGC TCC ACC Asn Pro Gly Ala Cys His Leu Ala Ser Pro Thr Ser Pro Pro Ser Thr 95 100 105	2017
40	GTA GCG GAC GGC CAG TAT CTA CAT CCA GAC TTC TCG GAG TCG TTC ACT Val Ala Asp Ala Gln Tyr Leu His Pro Asp Phe Ser Glu Ser Phe Thr 110 115 120 125	2065
45	CGA CTA CCA CCC CCA GAT CTC GTC TCC TCT CCC GAC TCG ACA AAC TCG Arg Leu Pro Pro Pro Asp Leu Val Ser Ser Pro Asp Ser Thr Asn Ser 130 135 140	2113
50	CTA TTC GAC TCG TCC ACT ATC GGC GCA CTC CCC GCG CCA CGC CGT CTG Leu Phe Asp Ser Ser Thr Ile Gly Ala Leu Pro Ala Pro Arg Arg Leu 145 150 155	2161
55	TCG AGC CCA AAC CTT CTA GGC CAT GTC AAT GTC TTC CTC AAG TAC CTG Ser Thr Pro Asn Leu Leu Ala His Val Asn Val Phe Leu Lys Tyr Leu 160 165 170	2209
60	TTC CCG ATC ATG CCC GTC GTG AGA CAG GAC CAG CTG CAG CAG GAC TGC Phe Pro Ile Met Pro Val Val Arg Gln Asp Gln Leu Gln Gln Asp Cys 175 180 185	2257
65	CAC CAG CCG GAG CGC TTG TCT CCC CAA CGC TAC GCT TTC ATT GGC GCT His Gln Pro Glu Arg Leu Ser Pro Gln Arg Tyr Ala Phe Ile Ala Ala 190 195 200 205	2305
70	CTA TGC GCG GGC ACG CAC ATC CAA CTG AAG CTG GAC GGT GCA CCA CCG Leu Cys Ala Ala Thr His Ile Gln Leu Lys Leu Asp Gly Ala Ala Pro 210 215 220	2353
75	GGT CCC GAG GCG GCT TCC GCG CCA GGC AGC CTC GAC GGA CAT CCT ATG Gly Pro Glu Ala Ser Ala Arg Ala Ser Leu Asp Gly His Pro Met 225 230 235	2401
80	TTG TCG GGA GAA GAA CTC CTG GCT GAA GGC GTG CGC CCA AGA AAG GAA Leu Ser Gly Glu Glu Leu Leu Ala Glu Ala Val Arg Ala Arg Lys Glu 240 245 250	2449
85	TAC AAC GTG GTC GAC GAA ATT AAC ATG GAA AAC CTC CTA ACC TCC TTC	2497

39

	Tyr Asn Val Val Asp Glu Ile Asn Met Glu Asn Leu Leu Thr Ser Phe	
	255 260 265	
5	TTT CTC TTC GCC GCC TAC GGA AAC CTA GAC AGA CAG GAT CAG GGC TGG Phe Leu Phe Ala Ala Tyr Gly Asn Leu Asp Arg Gln Asp Gln Ala Trp 270 275 280 285	2545
10	TTC TAC CTA TGT CAG ACC ACG TOC ATG GTC TTC ACA CTA GGC CTA CAA Phe Tyr Leu Cys Gln Thr Thr Ser Met Val Phe Thr Leu Gly Leu Gln 290 295 300	2593
15	CGG GAA TCC ACA TAC TCG AAA CTA AGC GTC CAG GAA GCA GAA GAG AAA Arg Glu Ser Thr Tyr Ser Lys Leu Ser Val Glu Glu Ala Glu Glu Lys 305 310 315	2641
20	AGG AGA GTA TTC TGG CTC TTA TTC GTC ACA GAA AG GTAAGAAAAG Arg Arg Val Phe Trp Leu Leu Phe Val Thr Glu Arg 320 325	2686
25	AAAAAATCT ACTTTCCCAA TCACCACCAC GTACCAAAAA TAACACGAAA AACCG A GOC TAC GCA TTA CAA CAA GCA AAA CCA GTC ATG CTC GGC AAC TCC ATC Gly Tyr Ala Leu Gln Gln Ala Lys Pro Val Met Leu Arg Asn Ser Ile 330 335 340 345	2743
30	CAC AAA CCA CAG GTC CTG TGC TCA GAC GAC CCA ATC CTA GGC TAC GGC His Lys Pro Gln Val Leu Cys Ser Asp Asp Pro Ile Leu Ala Tyr Gly 350 355 360	2791
35	TTC ATC AAC CTC ATC AAC GTC TTC GAA AAG CTC AGC CCA AAT CTC TAC Phe Ile Asn Leu Ile Asn Val Phe Glu Lys Leu Ser Pro Asn Leu Tyr 365 370 375	2839
40	GAC TGG GTC TOC GCC GGC GGC AGC AGC GCA GAC GGC GAC CCC CCG CCT Asp Trp Val Ser Ala Gly Gly Ser Ser Ala Asp Gly Asp Pro Pro Pro 380 385 390	2887
45	ACT TCT TCT ATC CAA TOC AGT CTC GOC AAG CAA ATC TOC CTC GAG GGC Thr Ser Ser Ile Gln Ser Ser Leu Ala Lys Gln Ile Ser Leu Glu Gly 395 400 405	2935
50	GTC TOC GAG ATC CAG AAA GTA GAC ATC CTC ATC ACT CAG CAA TGG CTA Val Ser Glu Ile Gln Lys Val Asp Ile Leu Ile Thr Gln Gln Trp Leu 410 415 420 425	2983
55	CAA ACC ATG ATG TGG AAA CTC TOC ATG ACC CAC GTC ACA CAG CCC GGC Gln Thr Met Met Trp Lys Leu Ser Met Thr His Val Thr Gln Pro Gly 430 435 440	3031
60	TCT GGC GAT GAC GGC GTT CTC OCC TTC CAC CTG CCC GTG CTA GTC GGC Ser Arg Asp Asp Ala Val Leu Pro Phe His Leu Pro Val Leu Val Gly 445 450 455	3079
65	AAG GGC GTC ATG GGC GTC ATC GGC GGC GCA TOC CAA GGT GCT GTT GAC Lys Ala Val Met Gly Val Ile Ala Ala Ala Ser Gln Gly Ala Val Asp 460 465 470	3127
	GCT CAT GGT ATC GGA ATG GTAAGAAAG GACCTTAOCT CATCACACC Ala His Gly Ile Gly Met 475	3175
	TCCTCATCA GTCACTCCCC ATCATCTATA CCCCACATCT AACAAAAACC GCAG GAA Glu 480	3223
	CAA AAA CTC TAC GAC CTC GGC ACC TOC GTA GGC GAC GTC TOC GGC TOC Gln Lys Leu Tyr Asp Leu Gly Thr Ser Val Ala Asp Val Ser Arg Ser 485 490 495	3280
		3328

40

	CTA AGC ACA AAA GCC GGC CAC CAC CTC GGC GAA TGG ACC ATC GAC CCC Leu Ser Thr Lys Ala Ala His His Leu Ala Glu Ser Thr Ile Asp Pro 500 505 510	3376
5	CGA GAA CTC CTC TGG GGC ATT CTC ACA ACC CTA TOC CGA ATC GGC GGT Arg Glu Leu Leu Trp Gly Ile Leu Thr Thr Leu Ser Arg Ile Arg Gly 515 520 525	3424
10	TOC CAA TCA TAC CTC TTC CCA GCG CTC GTC GAG CAA AGT CGA GGC ATC Ser Gln Ser Tyr Leu Phe Pro Ala Leu Val Glu Gln Ser Arg Gly Ile 530 535 540	3472
15	ATC AGT TTC GAC TGT TGG CTT TOC ATC AGT GAC TTT CTG CCT TGG TTT Ile Ser Phe Asp Cys Ser Leu Ser Ile Ser Asp Phe Leu Pro Ser Phe 545 550 555 560	3520
20	GGT GGG CCG CCG GCT ATT ATG TGG CCG AGG GGT GAA TCT GGG TTT GAT Gly Gly Pro Pro Ala Ile Met Trp Arg Thr Gly Glu Ser Gly Phe Asp 565 570 575	3568
25	TTA TTG GGG ATC GCG GAT GAT TTG CAA GAG AGG GAG AAT GAG GGT GGG Leu Leu Gly Ile Ala Asp Asp Leu Gln Glu Arg Glu Asn Glu Gly Gly 580 585 590	3616
30	GAG GGG ATT GTG GTG GCT GCG GAG GAG ATT TGG TTT TGAGGGGGCT Glu Gly Ile Val Val Ala Gly Glu Ile Ser Phe 595 600	3662
35	CTTTCTTTT TCCTTTTGGG TGCTGTGTGT TGGGTGGTTC TGGGGGGGGG GGGGTGTATA TACGCTTGAC GATGTGCATT GCGATTGGGG TTCTTACTGG TATAATATAT GCATTTTIT GTATATAGTC CGCTGGAGAC GGTGCAATGA TGTGGGGATC AATCACTTCT TAGGACTGGG AGCACAGGT GTGGTTCCTC GGGTATCTCT GAGTATGAGA TTATATAGAA TCAGTTATAG ATCATTTATG TACATACCTT AAAGAAAGAT ATGCTTGGCA CCGGATATG ACAATAGAAA ACTGGTCTTC ATTCAGA	3722 3782 3842 3902 3962
40		3980

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 604 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met	Ser	His	Ser	Pro	Thr	Asp	Ile	Pro	Ser	Thr	Ser	Glu	Lys	Glu	Met
1				5				10				15			
Glu	Ser	Thr	Pro	Glu	Lys	Pro	Pro	Lys	Gln	Ala	Cys	Asp	Asn	Cys	Arg
			20				25					30			
Arg	Arg	Lys	Ile	Lys	Cys	Ser	Arg	Glu	Leu	Pro	Cys	Asp	Lys	Cys	Gln
		35				40					45				
Arg	Leu	Leu	Leu	Ser	Cys	Ser	Tyr	Ser	Asp	Val	Leu	Arg	Arg	Lys	Gly
	50				55					60					
Pro	Lys	Phe	Arg	Thr	Leu	Tyr	Pro	Leu	Ala	Pro	Ile	His	Pro	Leu	Ala
65				70			75						80		
Ser	Arg	Pro	Arg	Pro	Leu	Thr	Lys	Glu	Trp	Leu	Pro	Pro	Asn	Pro	Gly
			85			90							95		

41

Ala Cys His Leu Ala Ser Pro Thr Ser Pro Pro Ser Thr Val Ala Asp  
100 105 110

5 Ala Gln Tyr Leu His Pro Asp Phe Ser Glu Ser Phe Thr Arg Leu Pro  
115 120 125

Pro Pro Asp Leu Val Ser Ser Pro Asp Ser Thr Asn Ser Leu Phe Asp  
130 135 140

10 Ser Ser Thr Ile Gly Ala Leu Pro Ala Pro Arg Arg Leu Ser Thr Pro  
145 150 155 160

Asn Leu Leu Ala His Val Asn Val Phe Leu Lys Tyr Leu Phe Pro Ile  
165 170 175

15 Met Pro Val Val Arg Gln Asp Gln Leu Gln Gln Asp Cys His Gln Pro  
180 185 190

20 Glu Arg Leu Ser Pro Gln Arg Tyr Ala Phe Ile Ala Ala Leu Cys Ala  
195 200 205

Ala Thr His Ile Gln Leu Lys Leu Asp Gly Ala Ala Pro Gly Pro Glu  
210 215 220

25 Ala Ala Ser Ala Arg Ala Ser Leu Asp Gly His Pro Met Leu Ser Gly  
225 230 235 240

Glu Glu Leu Leu Ala Glu Ala Val Arg Ala Arg Lys Glu Tyr Asn Val  
245 250 255

30 Val Asp Glu Ile Asn Met Glu Asn Leu Leu Thr Ser Phe Phe Leu Phe  
260 265 270

35 Ala Ala Tyr Gly Asn Leu Asp Arg Gln Asp Gln Ala Trp Phe Tyr Leu  
275 280 285

Cys Gln Thr Thr Ser Met Val Phe Thr Leu Gly Leu Gln Arg Glu Ser  
290 295 300

40 Thr Tyr Ser Lys Leu Ser Val Glu Glu Ala Glu Glu Lys Arg Arg Val  
305 310 315 320

Phe Trp Leu Leu Phe Val Thr Glu Arg Gly Tyr Ala Leu Gln Gln Ala  
325 330 335

45 Lys Pro Val Met Leu Arg Asn Ser Ile His Lys Pro Gln Val Leu Cys  
340 345 350

50 Ser Asp Asp Pro Ile Leu Ala Tyr Gly Phe Ile Asn Leu Ile Asn Val  
355 360 365

Phe Glu Lys Leu Ser Pro Asn Leu Tyr Asp Trp Val Ser Ala Gly Gly  
370 375 380

55 Ser Ser Ala Asp Gly Asp Pro Pro Pro Thr Ser Ser Ile Gln Ser Ser  
385 390 395 400

Leu Ala Lys Gln Ile Ser Leu Glu Gly Val Ser Glu Ile Gln Lys Val  
405 410 415

60 Asp Ile Leu Ile Thr Gln Gln Trp Leu Gln Thr Met Met Trp Lys Leu  
420 425 430

65 Ser Met Thr His Val Thr Gln Pro Gly Ser Arg Asp Asp Ala Val Leu  
435 440 445

Pro Phe His Leu Pro Val Leu Val Gly Lys Ala Val Met Gly Val Ile  
450 455 460

42

Ala Ala Ala Ser Gln Gly Ala Val Asp Ala His Gly Ile Gly Met Glu  
 465 470 475 480  
 5 Gln Lys Leu Tyr Asp Leu Gly Thr Ser Val Ala Asp Val Ser Arg Ser  
 485 490 495  
 Leu Ser Thr Lys Ala Ala His His Leu Ala Glu Ser Thr Ile Asp Pro  
 500 505 510  
 10 Arg Glu Leu Leu Trp Gly Ile Leu Thr Thr Leu Ser Arg Ile Arg Gly  
 515 520 525  
 Ser Gln Ser Tyr Leu Phe Pro Ala Leu Val Glu Gln Ser Arg Gly Ile  
 530 535 540  
 15 Ile Ser Phe Asp Cys Ser Leu Ser Ile Ser Asp Phe Leu Pro Ser Phe  
 545 550 555 560  
 20 Gly Gly Pro Pro Ala Ile Met Trp Arg Thr Gly Glu Ser Gly Phe Asp  
 565 570 575  
 Leu Leu Gly Ile Ala Asp Asp Leu Gln Glu Arg Glu Asn Glu Gly Gly  
 580 585 590  
 25 Glu Gly Ile Val Val Ala Gly Glu Glu Ile Ser Phe  
 595 600

## (2) INFORMATION FOR SEQ ID NO: 4:

30

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: primer 4650

(iii) HYPOTHETICAL: YES

40

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

45 CTTGCATGCC GCCAGGACCG AGCAAG 26

## (2) INFORMATION FOR SEQ ID NO: 5:

50

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55

(ii) MOLECULE TYPE: primer 4651

(iii) HYPOTHETICAL: YES

60

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

65 CTTGGATCCT CTGTGTTAGC TTATAG 26

## (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs

43

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: primer

(iii) HYPOTHETICAL: YES

10 (iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCCCAAGCTT CGCCGCTCTGC GCTGCTGCCG30

15 (2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29x base pairs  
20 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: primer

(iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CGGAATTCAT CAACCTCATC AACGTCCTC 29

35 (2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
40 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

45 (iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

50 CGGAATTCAT CGGCGAGATA GTATCCTAT 29

(2) INFORMATION FOR SEQ ID NO: 9:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 41 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
60 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(iii) HYPOTHETICAL: YES

65 (iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

WO 98/01470

PCT/DK97/00305

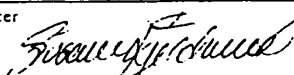
44

TTTGTAAAGC TTTTTTTT TTTTTTTT TTTTTTTT T

41

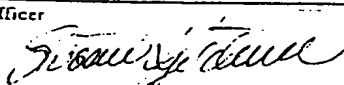
## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>11</u> , line <u>11-13</u>	
B. IDENTIFICATION OF DEPOSIT <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <b>DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH</b>	
Address of depositary institution (including postal code and country)  <b>Mascheroder Weg 1b, D-38124 Braunschweig, GERMANY</b>	
Date of deposit <b>1996-05-10</b>	Accession Number <b>DSM 10671</b>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<p>Until the publication of the mention of grant of a European patent or, where applicable, for twenty years from the date of filing if the application has been refused, withdrawn or deemed withdrawn, a sample of the deposited microorganism is only to be provided to an independent expert nominated by the person requesting the sample (cf. Rule 28(4) EPC). And as far as Australia is concerned, the expert option is likewise requested, reference being had to Regulation 3.25 of Australia Statutory Rules 1991 No 71. Also, for Canada we request that only an independent expert nominated by the Commissioner is authorized to have access to a sample of the microorganism deposited.</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer </p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description</b> on page <u>11</u> , line <u>31-34</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <b>DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH</b>	
Address of depositary institution (including postal code and country)  <b>Mascheroder Weg 1b, D-38124 Braunschweig, GERMANY</b>	
Date of deposit <b>1996-05-10</b>	Accession Number <b>DSM 10666</b>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Until the publication of the mention of grant of a European patent or, where applicable, for twenty years from the date of filing if the application has been refused, withdrawn or deemed withdrawn, a sample of the deposited microorganism is only to be provided to an independent expert nominated by the person requesting the sample (cf. Rule 28(4) EPC). And as far as Australia is concerned, the expert option is likewise requested, reference being had to Regulation 3.25 of Australia Statutory Rules 1991 No 71. Also, for Canada we request that only an independent expert nominated by the Commissioner is authorized to have access to a sample of the microorganism deposited.	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<b>For receiving Office use only</b> <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer 	<b>For International Bureau use only</b> <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

## CLAIMS

1. A transcription factor regulating the expression of an  $\alpha$ -amylase promoter in filamentous fungus.
- 5 2. The factor of claim 1 originating from a fungus of the genus *Aspergillus*, *Trichoderma*, *Penicillium*, *Fusarium*, *Humicola*, etc.
- 10 3. The factor of claim 2 originating from the species *A. oryzae*, *A. niger*, *A. awamori*, especially *A. oryzae* IFO4177.
4. The factor of claim 3 having an amino acid sequence comprising one or more fragments of the amino acid sequence depicted as SEQ. ID. No 3.
- 15 5. A DNA construct having a DNA sequence coding for the factor of any of the claims 1 to 4.
- 20 6. The DNA sequence of claim 5 having a DNA sequence comprising one fragment or a combination of fragments of the DNA sequence depicted as SEQ ID NO:1.
7. A DNA construct comprising a DNA sequence encoding a transcription factor exhibiting activity in regulating the expression of an  $\alpha$ -amylase promoter in a filamentous fungus, which DNA sequence comprises
  - a) the transcription factor encoding part of the DNA sequence cloned into plasmid pToC320 present in *E. coli* ToC1058, DSM 30 10666, or
  - b) an analogue of the DNA sequence defined in a), which
    - i) is at least 60% homologous with the DNA sequence defined in a), or
    - ii) hybridizes with the same nucleotide probe as the DNA sequence defined in a), or
    - 35 iii) encodes a transcription factor which is at least 50% homologous with the transcription factor

- encoded by a DNA sequence comprising the DNA sequence defined in a), or
- iv) encodes a transcription factor which is immunologically reactive with an antibody raised against the purified transcription factor encoded by the DNA sequence defined in a), or
- v) complements the mutation in ToC879, i.e. makes ToC879 able to grow on cyclodextrin and produce lipase when transformed with said DNA sequence.
8. The DNA construct according to any of the claims 5 to 7, in which said DNA sequence is obtainable from a filamentous fungus.
9. The DNA construct according to claim 8, in which said filamentous fungus belongs to any of the groups Phycomycetes, Zygomycetes, Ascomycetes, Basidiomycetes and fungi imperfecti, including Hyphomycetes such as the genera *Aspergillus*, *Penicillium*, *Trichoderma*, *Fusarium* and *Humicola*, in particular a strain from *Aspergillus* sp., and especially from *A. oryzae*.
10. The DNA construct according to claim 9, in which said DNA sequence is isolated from or produced on the basis of a DNA library of an *Aspergillus oryzae* strain.
11. The DNA construct according to claim 5 to 8, in which said DNA sequence is obtainable from a yeast strain, especially of, *Saccharomyces*.
12. The DNA construct according to claim 7, in which the DNA sequence is isolated from *Escherichia coli* DSM 10666.
13. A recombinant expression vector comprising a DNA construct according to any of claims 5 to 12.
14. A cell comprising a DNA construct according to any of claims 5 to 12, or a recombinant expression vector according to claim 13.

15. The cell according to claim 14, which is a eukaryotic cell, in particular a fungal cell, such as a yeast cell or a filamentous fungal cell.

5

16. The cell according to claim 15, which is a strain of *Aspergillus* sp, in particular a strain of *A. niger* or *A. oryzae*.

10

17. The cell according to claim 15, which is a strain of *Trichoderma* sp., in particular *T. reesei*.

18. The cell according to claim 15, which is a strain of *Saccharomyces*, in particular a strain of *S. cerevisiae*.

15

19. A method of producing a polypeptide of interest comprising growing a cell of any of the claims 14 to 18 under conditions conducive to the production of said factor and said polypeptide of interest, and recovering said polypeptide of  
20 interest.

20. The method of claim 19, wherein said fungus is a fungus of the genus *Aspergillus*, *Trichoderma*, *Penicillium*, *Fusarium* or *Humicola*.

25

21. The method of claim 20, wherein said cell is of the species *A. oryzae*, *A. niger*, or *A. awamori*.

22. The method of claim 19, 20, or 21, wherein said  
30 polypeptide of interest is a medicinal polypeptide.

23. The method of claim 22, wherein said medicinal polypeptide is a growth hormone, insulin, or a blood clotting factor.

35 24. The method of claim 19, 20, or 21, wherein said polypeptide is an industrial enzyme.

50

25. The method of claim 24, wherein said industrial enzyme is a carbonyl hydrolase, carbohydrase, protease, lipase, amylase, cellulase, oxido reductase, glucoamylase, or esterase.
- 5 26. Use of a factor of any of the claims 1 to 4 for enhancing the expression of a polypeptide of interest in a filamentous fungus.
- 10 27. The use of claim 26, wherein said factor is the factor of claim 4.
- 15 28. The use of claim 27, wherein said fungus is a fungus of the genus *Aspergillus*, *Penicillium*, *Trichoderma*, *Fusarium* and *Humicola*, in particular a strain from *Aspergillus* sp., and especially from *A. oryzae* sp.
- 20 29. The use of claim 28, wherein said fungus is of the species *A. oryzae*, *A. niger*, *A. awamori*, *T. reesei*, or *T. harzianum*.
30. The use of any of the claims 26 to 29, wherein said polypeptide of interest is a medicinal polypeptide.
- 25 31. The use of claim 30, wherein said medicinal polypeptide is a growth hormone, insulin, or blood clotting factor.
32. The use of any of the claims 26 to 29, wherein said polypeptide is an industrial enzyme.
- 30 33. The use of claim 32, wherein said industrial enzyme is a carbonyl hydrolase, carbohydrase, protease, lipase, amylase, cellulase, oxido reductase, glucoamylase, or esterase.

1/8

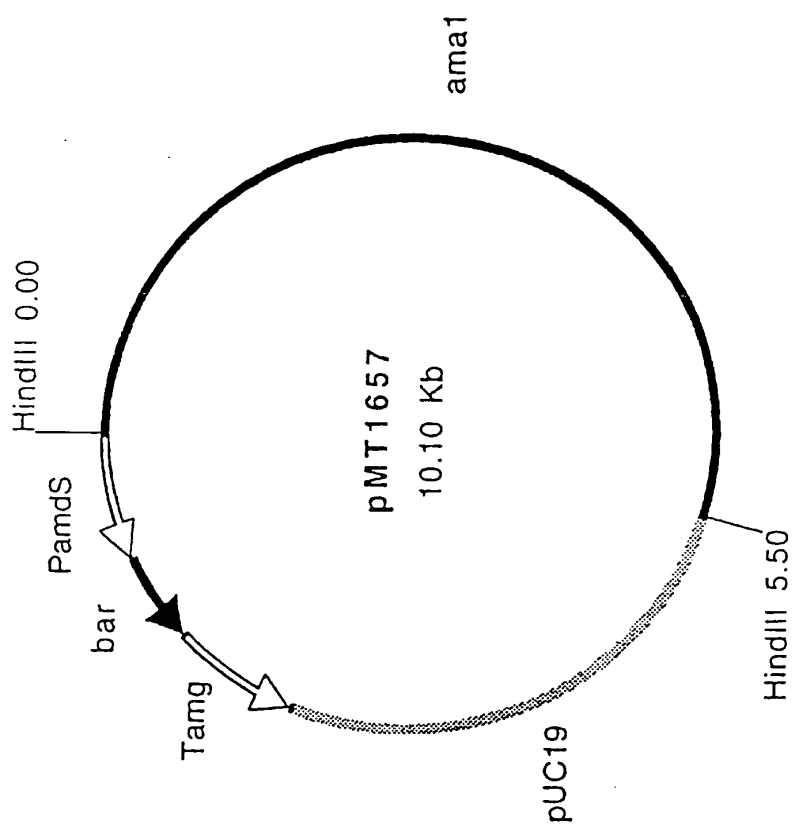


Fig. 1

2/8

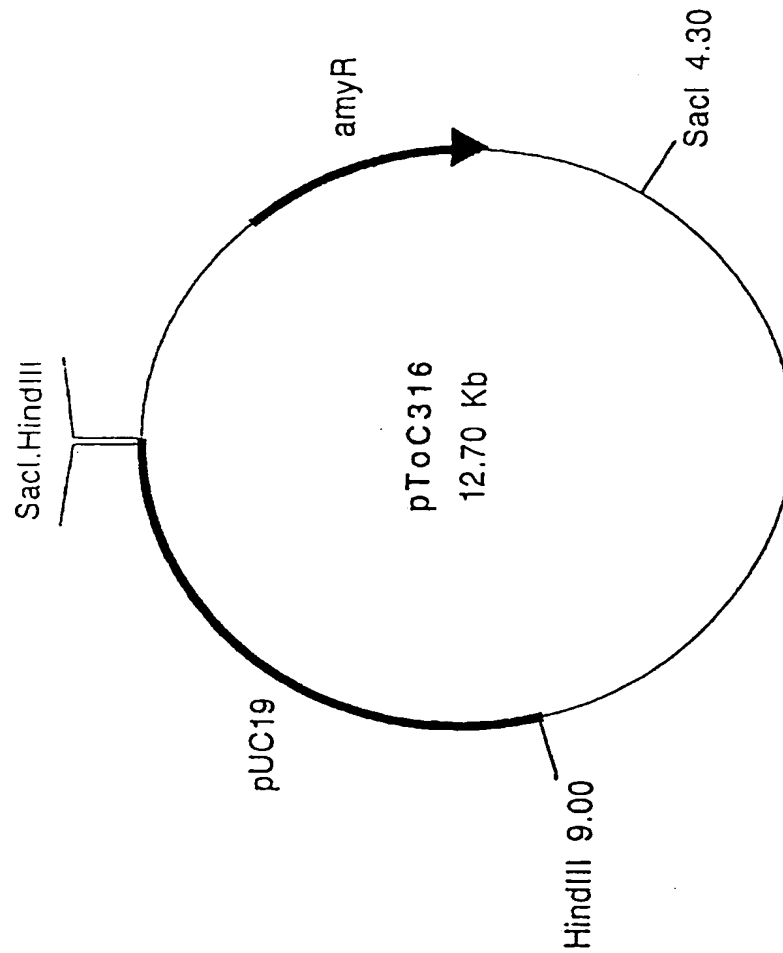


Fig. 2

3/8

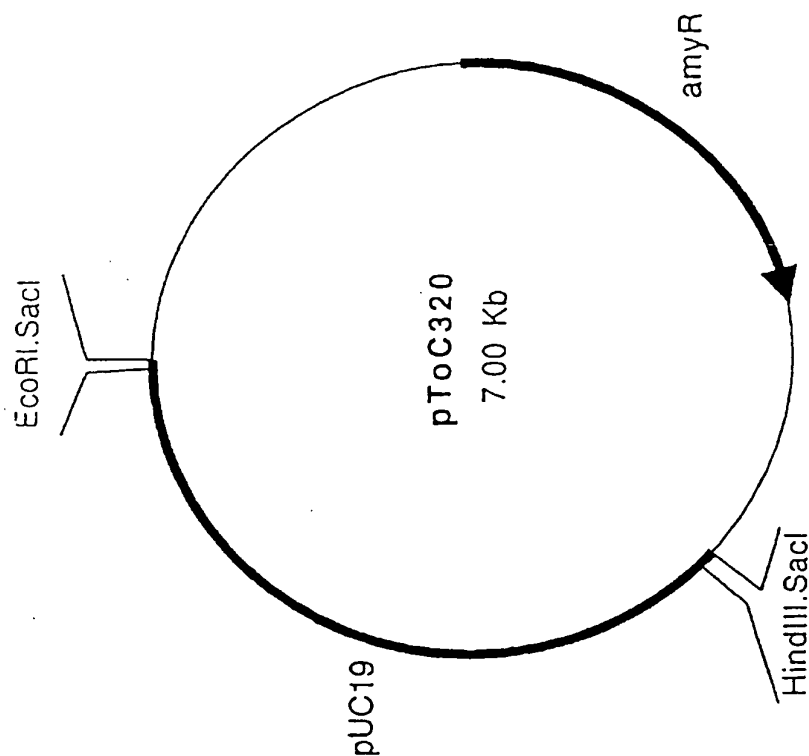


Fig. 3

4/8

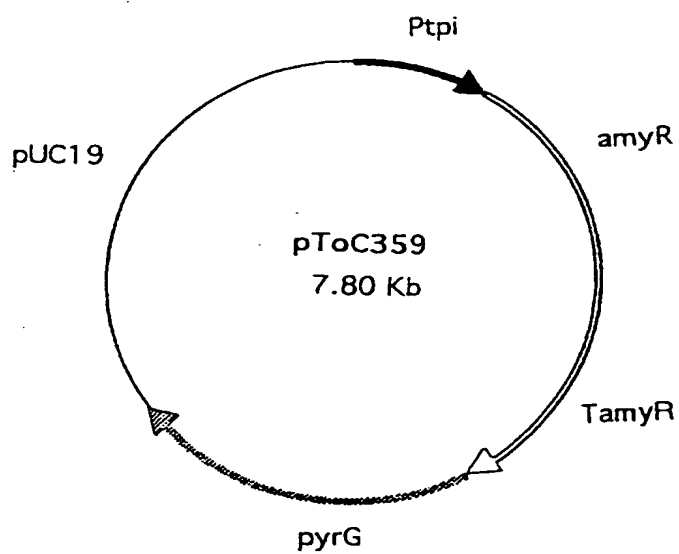
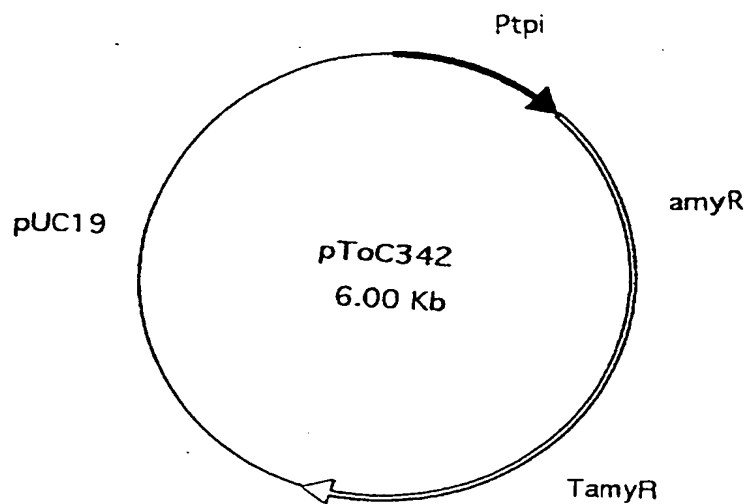


Fig. 4

5/8

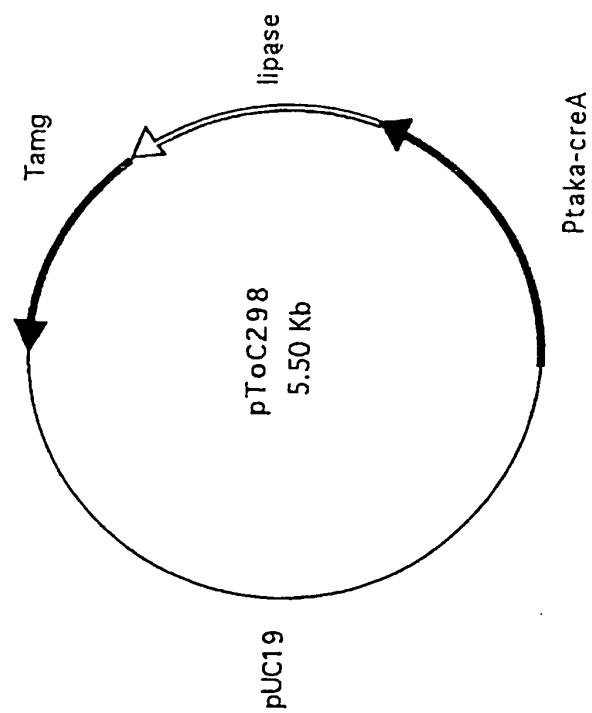


Fig. 5

6/8

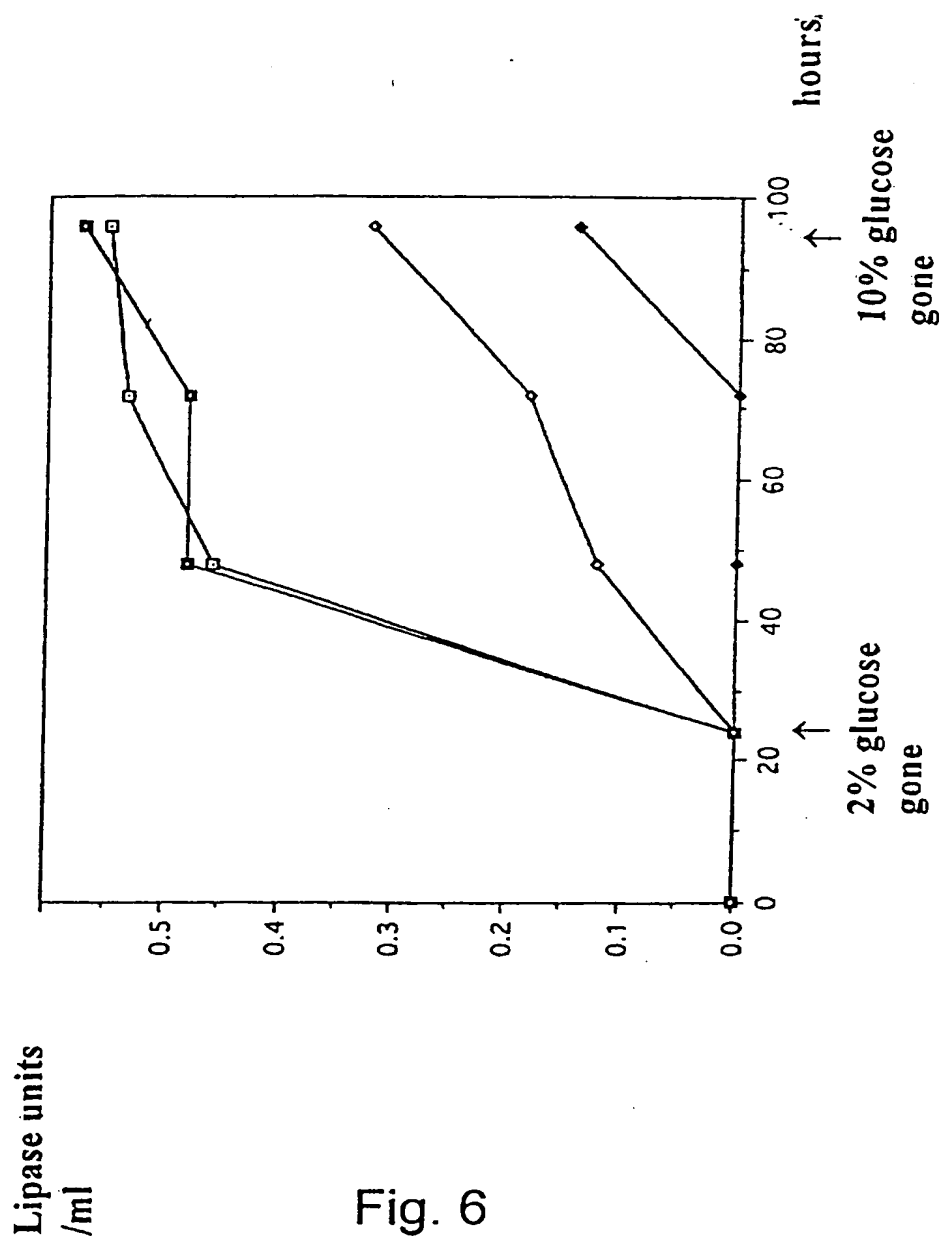


Fig. 6

7/8

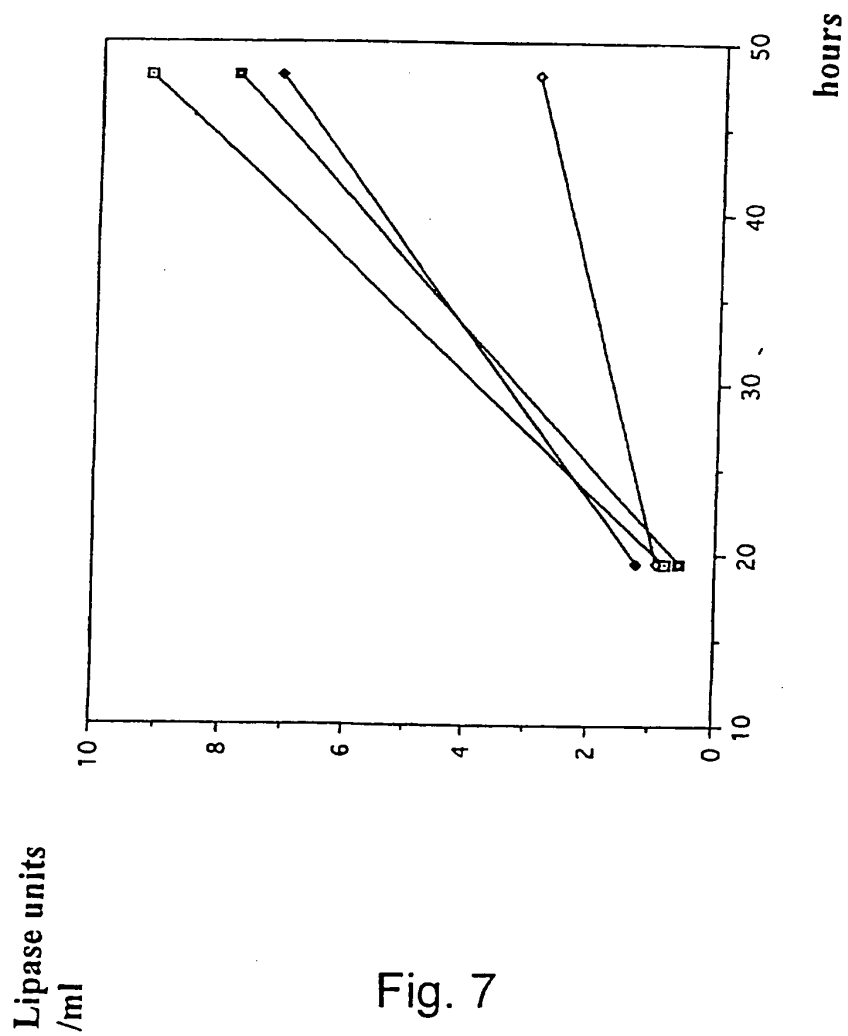


Fig. 7

8/8

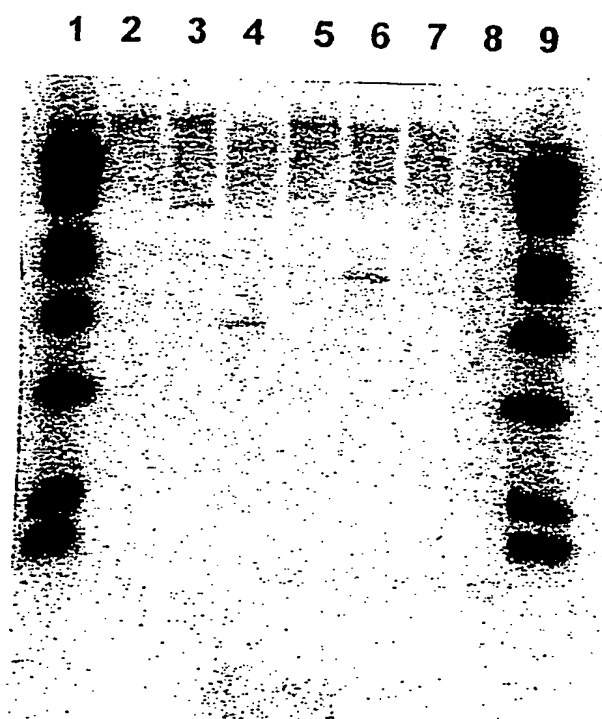


Fig. 8

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/DK 97/00305

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC6: C07K 14/38, C12N 15/80, C12N 1/15 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC6: C12N, C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,DK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI, EDOC, MEDLINE, BIOSIS, DBA, SCISEARCH GENBANK/SWISSPROT/EMBL/DBJ		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Chemical Abstracts, Volume 121, No 12, 19 Sept 1994 (19.09.94), (Columbus, Ohio, USA), Verdoes, Jan C. et al, "The effect of multiple copies of the upstream region on expression of the Aspergillus niger glucoamylase en coding gene", page 272, THE ABSTRACT No 150449j, Gene 1994, 145 (2), 179-187	1-3,5,8-10, 13-18
A		4,6-7,11-12, 19-33
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"B" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
Date of the actual completion of the international search		Date of mailing of the international search report
29 October 1997		07-11-1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer  Patrick Andersson Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 97/00305

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>Dialog Information Services, File 34, SciSearch, Dialog accession no. 13944964, Verdoes JC et al: "Molecular-Genetic Strain Improvement for the Overproduction of Fungal Proteins by Filamentous Fungi", Applied Microbiology and Biotechnology, 1995, V43, N2 (May-Jun), p 195-205</p> <p>---</p>	1-33
A	<p>Dialog Information Service, file 154, Medline, Dialog accession no. 07510263, Medline accession no. 93204901, Nagata O. et al: "Aspergillus nidulans nuclear proteins bind to a CCAAT element and the adjacent upstream sequence in the promoter region of the starch-inducible Taka-amylase A gene", Mol Gen Genet (GERMANY) Feb 1993, 237 (1-2) p251-60</p> <p>---</p> <p>-----</p>	1-33

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 97/00305

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 1, 7 and related claims  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
see next page
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees. - - - - -

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1992)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 97/00305

The wording of claim 1 "A transcription factor regulating the expression of an alpha amylase promoter in filamentous fungus" is not clear as promoters are not expressed i.e. the claim does not fulfill the prescribed requirements of a claim see Art 6 and Art 17(2)(a)(ii). The claim has been interpreted as " A transcription factor regulating alpha amylase promoter initiated expression in filamentous fungus"

**THIS PAGE BLANK (USPTO)**